OIPE IAP85
SEP 2 9 2005
PATENT & TRADEMARK OFFICE

Fig. 1a-f

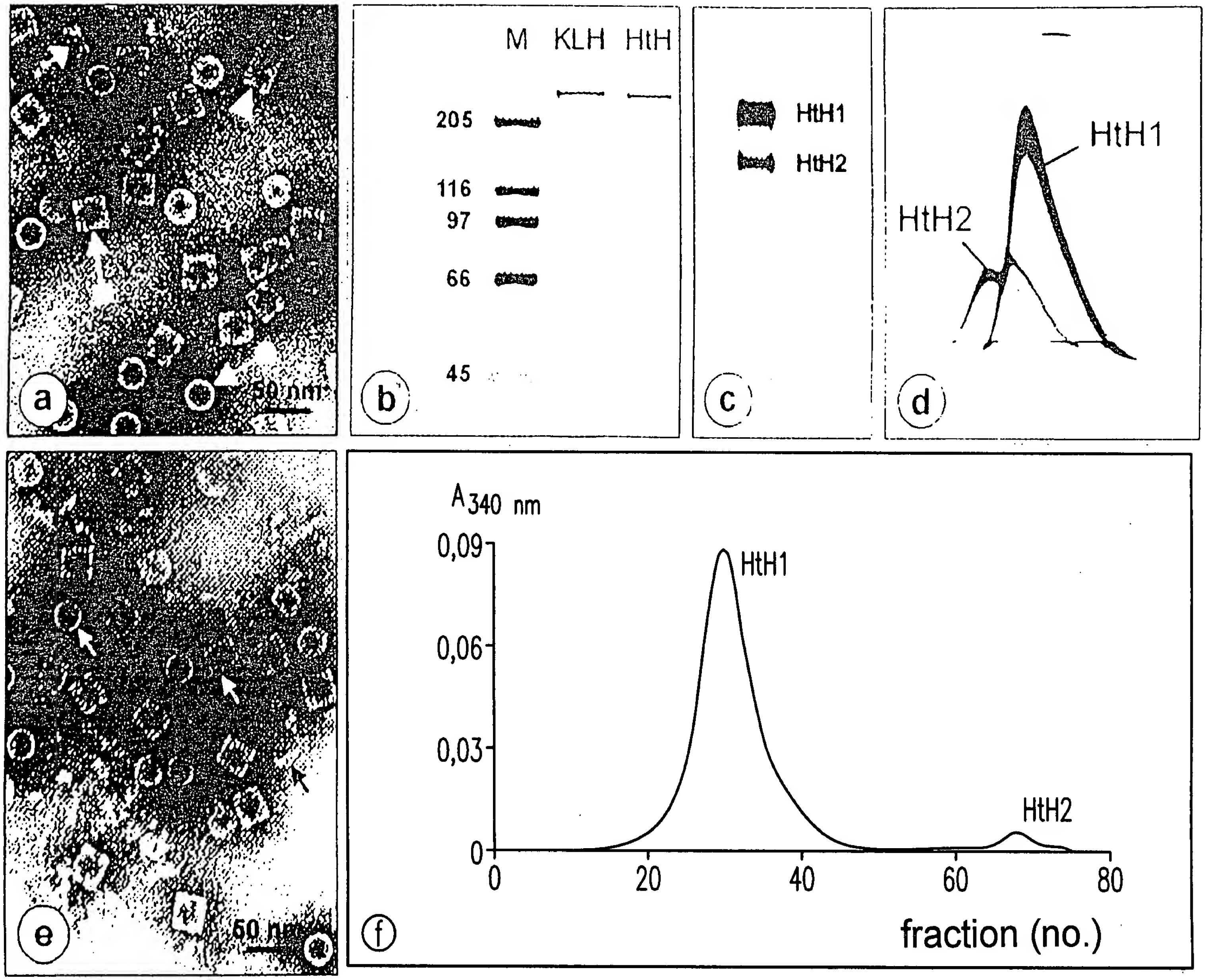
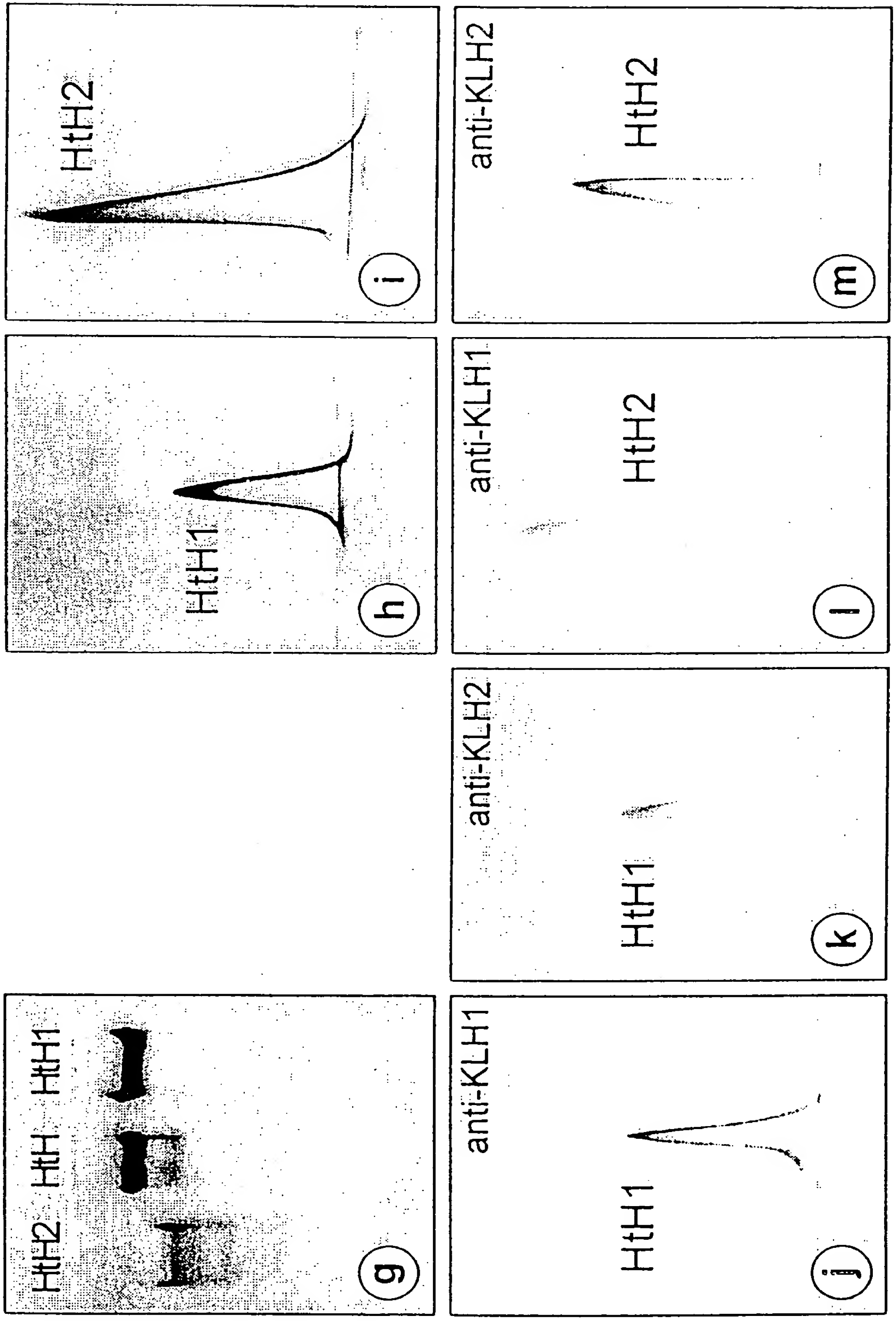


Fig. 1g-m


HtH2
HtH
HtH1
g
HtH1
h
HtH2
i
anti-KLH1
HtH1
j
anti-KLH2
HtH1
k
anti-KLH1
HtH2
l
anti-KLH2
HtH2
m

REPLACEMENT SHEET NO. 3

Fig. 2a-h

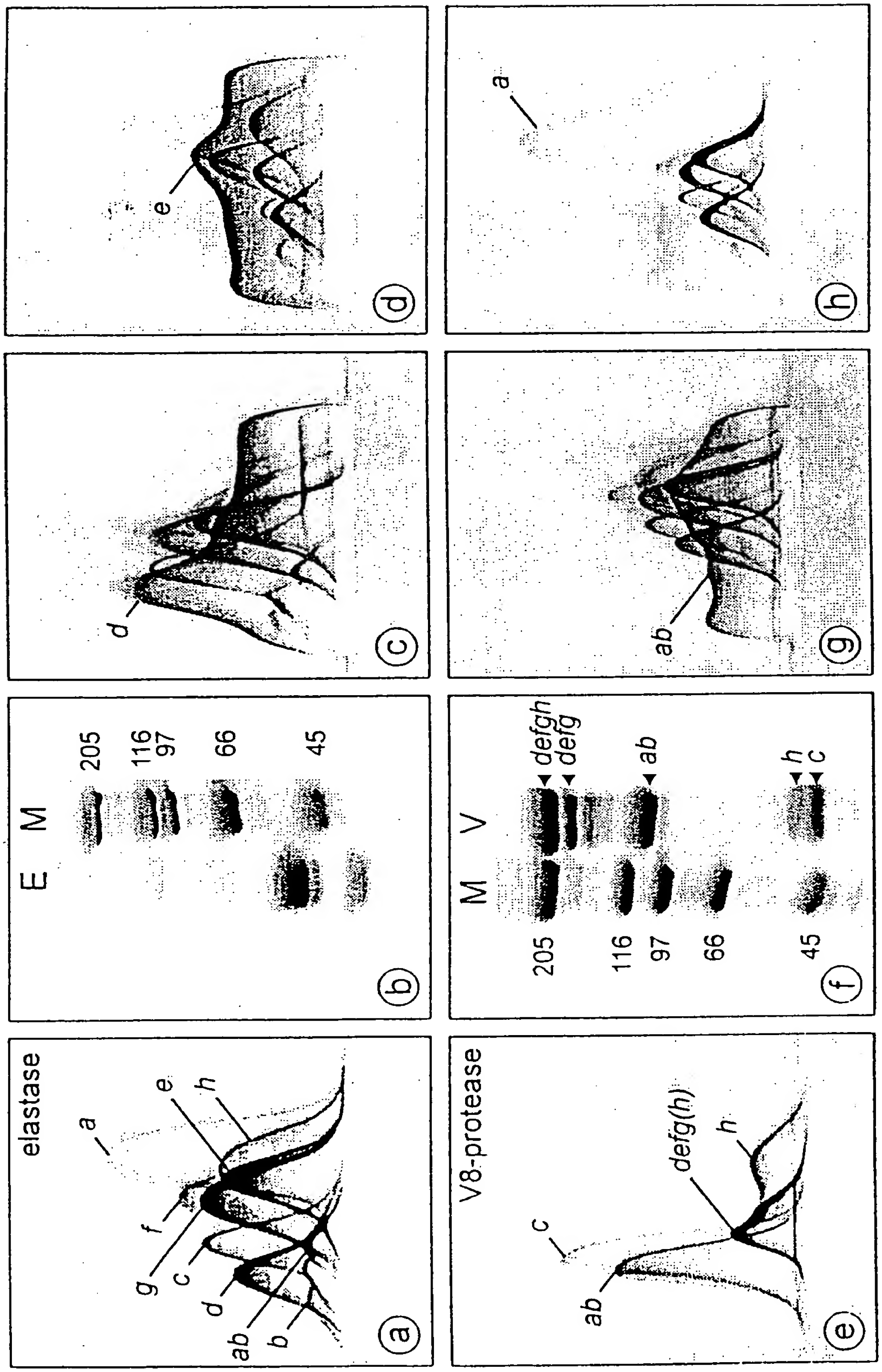


Fig. 2i-p

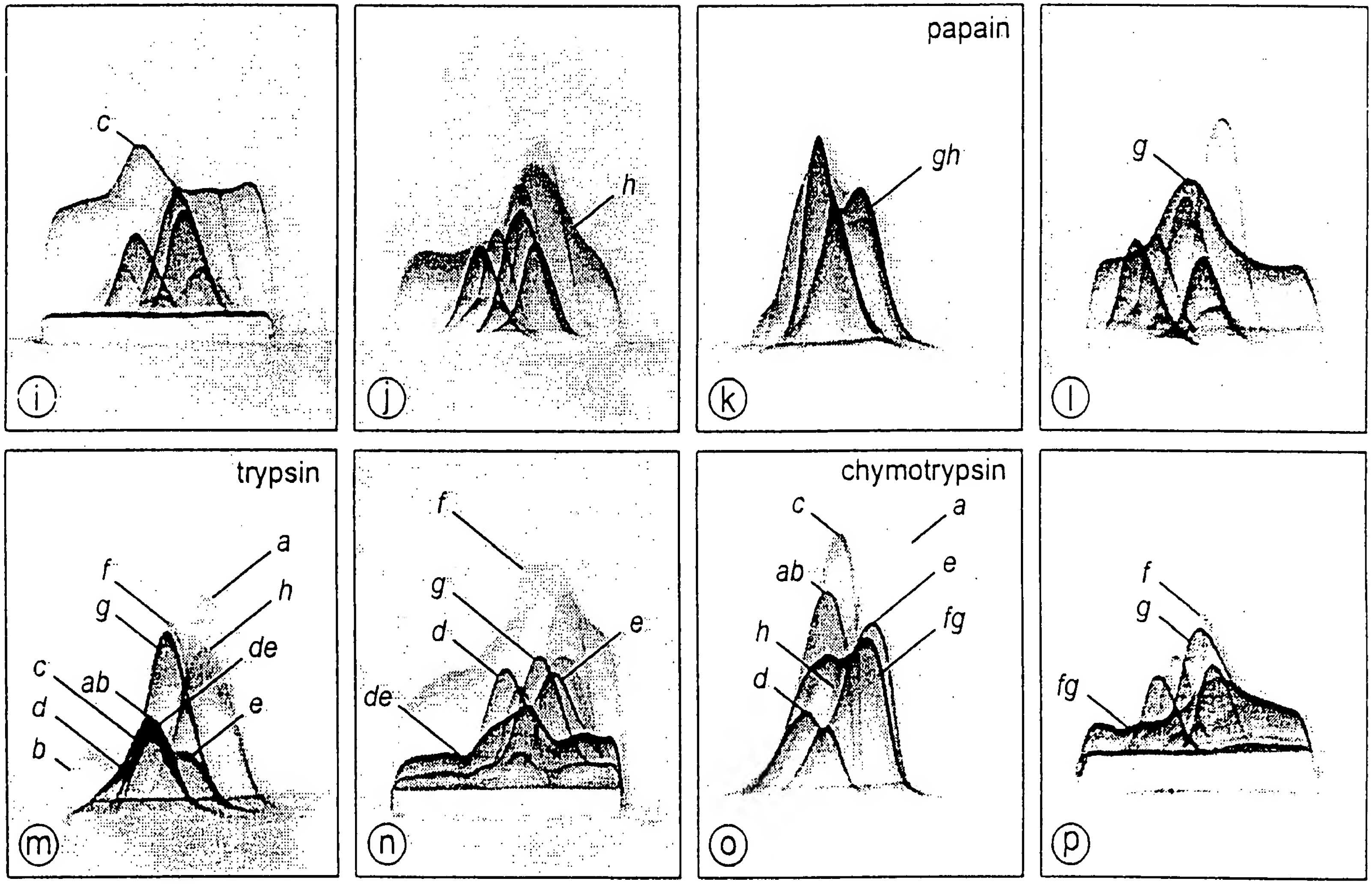

c
ⓘ
h
ⓙ
papain
gh
ⓚ
g
ⓛ
trypsin
a
f
h
g
de
c
ab
d
e
b
ⓜ
f
g
d
e
de
ⓝ
chymotrypsin
c
a
ab
e
h
fg
d
ⓞ
f
g
fg
ⓟ

Fig. 3a-c

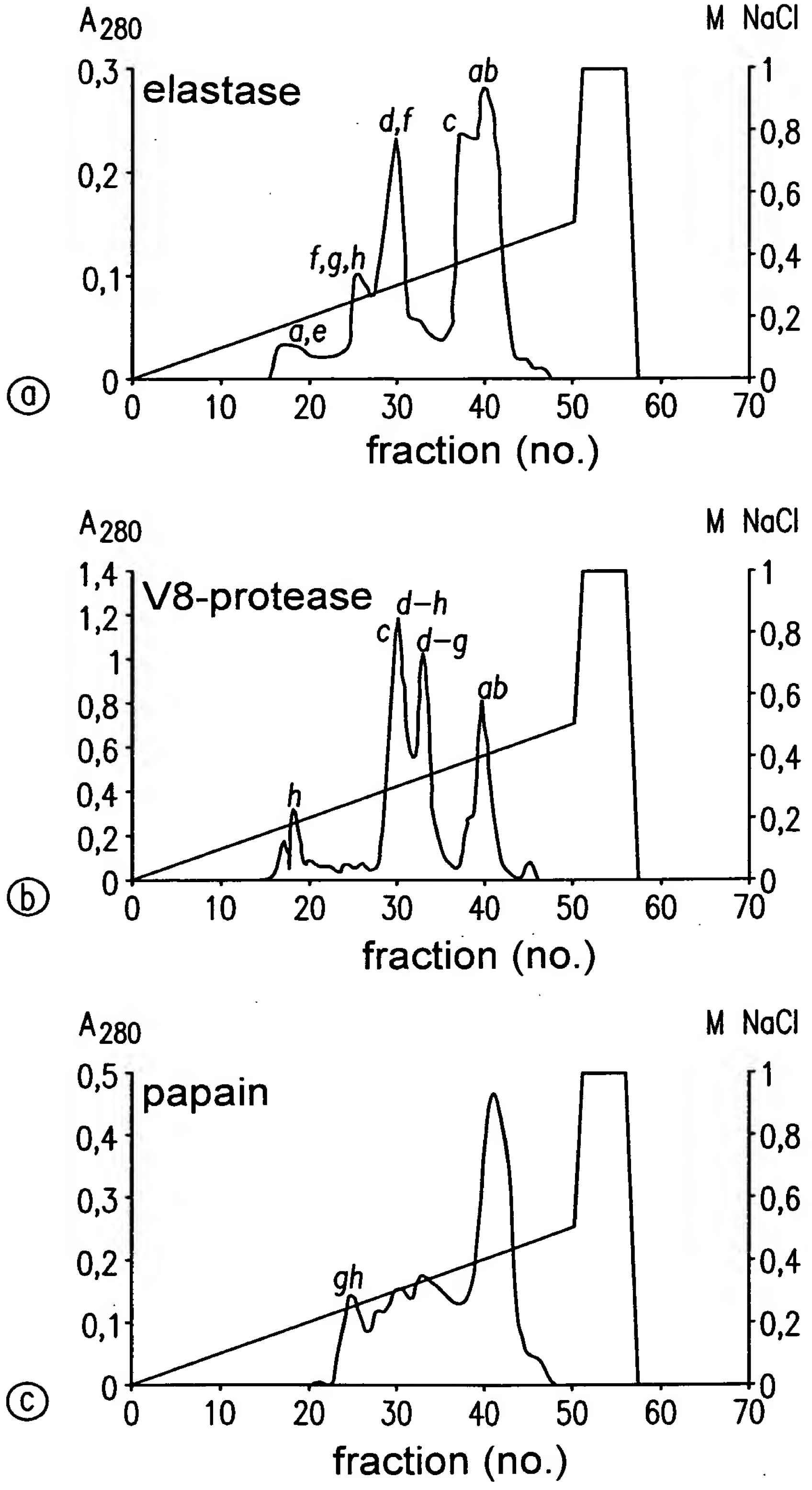

A280
M NaCl
elastase
0,3
0,2
0,1
0
ab
d,f
c
f,g,h
a,e
1
0,8
0,6
0,4
0,2
0
ⓐ
0 10 20 30 40 50 60 70
fraction (no.)
A280
M NaCl
V8-protease
1,4
1,2
1
0,8
0,6
0,4
0,2
0
d–h
c
d–g
ab
h
ⓑ
0 10 20 30 40 50 60 70
fraction (no.)
A280
M NaCl
papain
0,5
0,4
0,3
0,2
0,1
0
gh
ⓒ
0 10 20 30 40 50 60 70
fraction (no.)

Fig. 3d-e

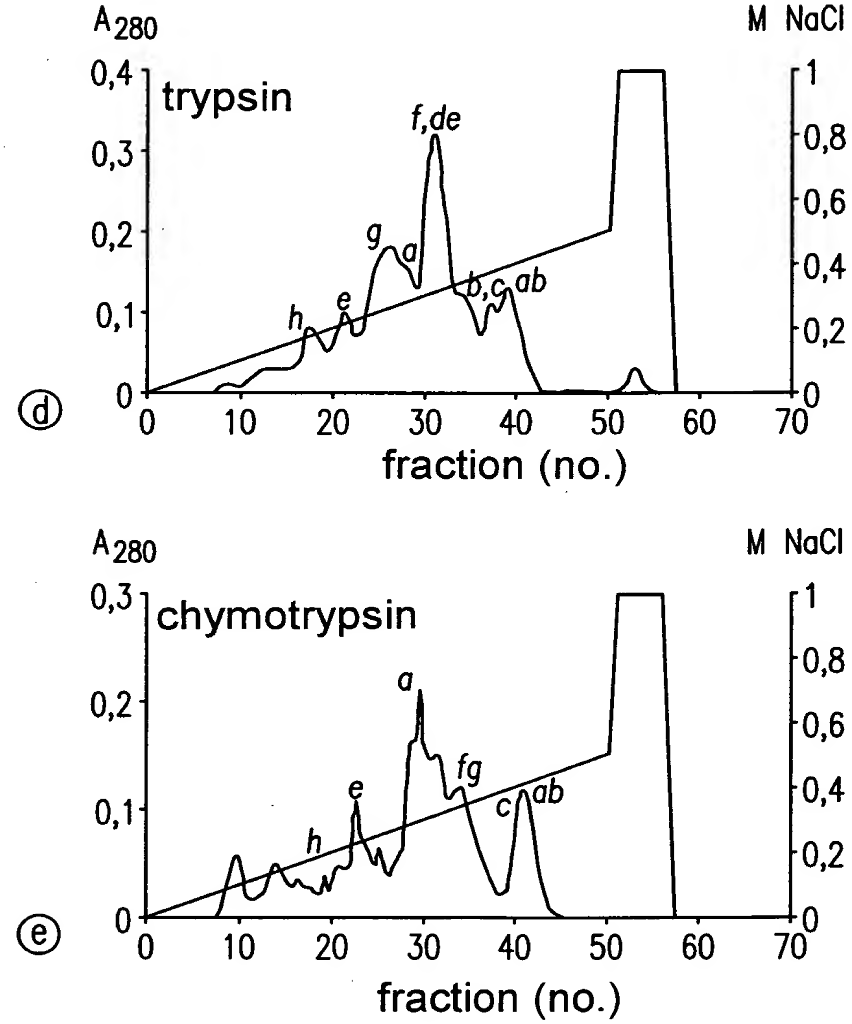

A280
M NaCl
0,4
0,3
0,2
0,1
0
trypsin
f,de
g
a
h
e
b,c
ab
1
0,8
0,6
0,4
0,2
0
ⓓ
0
10
20
30
40
50
60
70
fraction (no.)
A280
M NaCl
0,3
0,2
0,1
0
chymotrypsin
a
fg
e
h
c
ab
1
0,8
0,6
0,4
0,2
0
ⓔ
0
10
20
30
40
50
60
70
fraction (no.)

Figure 4

HtH1 cDNA sequence and intron structure

Domain a

GGCTTGTTCAGTTTCTACTCGTCGCCCTTGTGGCGGGGGCTGGAGCAGACAACGTCGTCAG
AAAGGACGTGAGTCACCTCACGGATGACGAGGTGCAAGCTCTCCACGGCGCCCTCCATGAC
GTCACTGCATCTACAGGGCCTCTGAGTTTCGAAGACATAACATCTTACCATGCCGCACCAG
CGTCGTGTGACTACAAGGGACGGAAGATCGCCTGCTGTGTCCACGGTATGCCCAGTTTCCC
CTTCTGGCACAGGGCATATGTCGTCCAAGCCGAGCGGGCACTGTTGTCCAAACGGAAGACT
GTCGGAATGCCTTACTGGGACTGGACGCAAACGCTGACTCACTTACCATCTCTTGTGACTG
AACCCATCTACATTGACAGTAAAGGTGGAAAGGCTCAAACCAACTACTGGTACCGCGGCGA
GATAGCGTTCATCAATAAGAAGACTGCGCGAGCTGTAGATGATCGCCTATTCGAGAAGGTG
GAGCCTGGTCACTACACACATCTTATGGAGACTGTCCTCGACGCTCTCGAACAGGACGAAT
TCTGTAAATTTGAAATCCAGTTCGAGTTGGCTCATAATGCTATCCATTACTTGGTTGGCGG
TAAATTTGAATATTCAATGTCAAACTTGGAATACACCTCCTACGACCCCATCTTCTTCCTC
CACCACTCCAACGTTGACCGCCTCTTCGCCATCTGGCAGCGTCTTCAGGAACTGCGAGGAA
AGAATCCCAATGCAATGGACTGTGCACATGAACTCGCTCACCAGCAACTCCAACCCTTCAA
CAGGGACAGCAATCCAGTCCAGCTCACAAAGGACCACTCGACACCTGCTGACCTCTTTGAT
TACAAACAACTTGGATACAGCTACGACAGCTTAAACCTGAATGGAATGACGCCAGAACAGC
TGAAAACAGAACTAGACGAACGCCACTCCAAAGAAGGTGCGTTTGCAAGCTTCCGACTCAG
TGGCTTTGGGGGTTCTGCCAACGTTGTTGTCTATGCATGTGTCCCTGATGATGATCCACGC
AGTGATGACTACTGCGAGAAAGCAGGCGACTTCTTCATTCTTGGGGGTCAAAGCGAAATGC
CGTGGAGATTCTACAGACCCTTCTTCTATGATGTAACTGAAGCGGTACATCACCTTGGAGT
CCCGCTAAGTGGCCACTACTATGTGAAAACAGAACTCTTCAGCGTGAATGGCACAGCACTT
TCACCTGATCTTCTTCCTCAACCAACTGTTGCCTACCGACCTGGGAAAG

Domain b

GTCACCTTGACCCACCTGTGCATCATCGCCACGATGACGATCTTATTGTTCGAAAAAATAT
AGATCATTTGACTCGTGAAGAGGAATACGAGCTAAGGATGGCTCTGGAGAGATTCCAGGCC
GACACATCCGTTGATGGGTACCAGGCTACAGTAGAGTACCATGGCCTTCCTGCTCGTTGTC
CACGACCAGATGCAAAAGTCAGGTTCGCCTGTTGTATGCATGGCATGGCATCCTTCCCTCA
CTGGCACCGGCTGTTCGTTACCCAGGTGGAAGATGCTCTTGTACGGCGTGGATCGCCTATC
GGTGTTCCTTATTGGGACTGGACAAAACCTATGACTCACCTTCCAGACTTGGCATCAAATG
AGACGTACGTAGACCCGTATGGACATACACATCATAATCCATTCTTCAATGCAAATATATC
TTTTGAGGAGGGACACCATCACACGAGCAGGATGATAGATTCGAAACTGTTTGCCCCAGTC
GCTTTTGGGGAGCATTCCCATCTGTTTGATGGAATCCTGTACGCATTTGAGCAGGAAGATT
TCTGCGACTTTGAGATTCAGTTTGAGTTAGTCCATAATTCTATTCATGCGTGGATAGGCGG
TTCCGAAGATTACTCCATGGCCACCCTGCATTACACAGCCTTTGACCCCATTTTCTACCTT
CATCATTCCAATGTCGATCGTCTATGGGCAATCTGGCAAGCTCTTCAAATCAGGAGACACA
AGCCATATCAAGCCCACTGTGCACAGTCTGTGGAACAGTTGCCAATGAAGCCATTTGCTTT
CCCATCACCTCTTAACAACAACGAGAAGACACATAGTCATTCAGTCCCGACTGACATTTAT
GACTACGAGGAAGTGCTGCACTACAGCTACGATGATCTAACGTTTGGTGGGATGAACCTTG
AAGAAATAGAAGAAGCTATACATCTCAGACAACAGCATGAACGAGTCTTCGCGGATTTCT
CCTTGCTGGAATAGGAACATCTGCACTTGTTGACATTTTCATAAATAAACCGGGGAACCAA
CCACTCAAAGCTGGAGATATTGCCATTCTTGGTGGTGCCAAGGAAATGCCTTGGGCGTTTG
ACCGCTTGTATAAGGTCGAAATAACTGACTCATTGAAGACACTTTCTCTCGATGTCGATGG
AGATTATGAAGTCACTTTTAAAATTCATGATATGCACGGAAACGCTCTTGATACGGACCTG
ATTCCACACGCAGCAGTTGTTTCTGAGCCAGCTCACC

Intron b/c

GTAAGTAAATTTACAAAATTTGGTGTTCTCTAACTATCCTAAGTATTCAATCGTTAGCGTG
TACCTATCTGCATAATGCAATACCCTGACTCCATATAAGTATAGTATATTTACTCTGGTCG
AAAACAAACAAATTGAAAACAAGAGTGGACGTGCTGTTATGATTTCTTTTTCATTCTTGGT
TCGTTGTGTAATGCCACAGCCAGCAATTCCAGATATATAGCGACGGTCTATGAATACTCCA
GTCTGGACCAGACAATCGTGTGGAATGGTTTAGGCACATTATATCAAATTCATTGTTGAAG
ATATGAGTTATGAGGTCACAATGTTGTCTTGTTACCCCGTGTCAGTAGTGACGTCATTTCA
TGACTGAAATCTCTTCAACGCCGTTTAGCAATAATAGGCTCAGTAGTATTCAACCAATTAC
AATCAGTAGAAAATTCTCTATACTATTCTTATGTTGCATCCTGATATCCCTATGCAAAAAT
TAGTCATCTAATATAATCATTTTCGATAAATACTTTGGGCAAACAAATCAATGTAACATCT
ATTTTCTTTCAG

Domain c

CTACCTTTGAGGATGAAAAGCACAGCTTACGAATCAGAAAAAATGTCGACAGCTTGACTCC
TGAAGAAACAAATGAACTGCGTAAAGCCCTGGAGCTTCTTGAAAATGATCATACTGCAGGT
GGATTCAATCAGCTTGGCGCCTTCCATGGAGAGCCTAAATGGTGCCCTAATCCTGAAGCGG
AGCACAAGGTTGCATGCTGTGTTCATGGCATGGCTGTTTTCCCTCATTGGCACAGGCTTCT
TGCTCTCCAGGCGGAGAATGCTCTTAGAAAGCATGGGTACAGTGGTGCTCTACCATACTGG
GATTGGACTCGCCCCCTTTCCCAACTTCCTGATCTGGTTAGTCATGAGCAGTATACAGATC
CTTCCGACCATCACGTGAAGCATAACCCGTGGTTCAATGGCCACATCGATACAGTAAATCA
GGATACCACCAGAAGCGTACGGGAGGATCTTTATCAACAACCTGAATTTGGACATTTCACG
GATATTGCTCAACAAGTCCTCTTAGCATTAGAACAAGATGACTTCTGTTCGTTTGAAGTGC
AGTATGAGATTTCCCATAATTTTATCCATGCACTTGTAGGAGGAACCGACGCTTATGGCAT
GGCATCGCTGAGATATACAGCATACGATCCAATCTTTTTCTTGCATCATTCAAACACCGAC
AGGATCTGGGCTATTTGGCAATCCCTGCAAAAATACAGAGGCAAACCGTACAACACTGCCA
ACTGCGCCATAGAATCTATGAGAAGGCCCCTGCAACCATTTGGACTAAGCAGTGCCATTAA
CCCTGACAGAATCACCAGAGAGCATGCTATCCCGTTTGATGTCTTCAACTATAGAGATAAC
CTTCATTACGTATATGATACCCTGGAATTTAATGGTTTGTCGATTTCACAACTTGATAGAG
AGCTGGAAAAAATCAAGAGTCACGAAAGAGTATTTGCTGGATTCTTGCTGTCGGGGATTAA
AAAATCTGCTCTTGTGAAATTCGAAGTTTGTACTCCACCTGATAATTGTCATAAAGCAGGG
GAGTTTTATCTACTCGGGGACGAAAACGAGATGGCTTGGGCCTATGACCGACTTTTCAAGT
ATGATATTACTCAGGTTCTGGAAGCAAACCATCTACACTTCTATGATCATCTCTTCATTCG
CTACGAAGTCTTTGATCTTAAAGGAGTGAGTTTGGGAACTGACCTGTTCCACACTGCAAAT
GTGGTACATGATTCCGGCACAG

Intron c/d

GTACGTGGATTTGATTACATAGCAATGCTATATGATTTCAGTAATTACAACCTCAAGTCAT
GTAGCCGTTTTAGATTGCATTACATCAAACAGCATTGGATTAAATTGGGGATTGTCCAGG
CCGCATTATGTTGCATTCCGAAAATAGTTTGTGTCCAGTGTCCACGTTTAAAATTAAACCA
TTTTAATCATATTAGGGATAATTTTAATAGATGTTATAGTGCTTTATTTCATATTGTTACA
GTGGACAGTCACCAAGGACATATTTTACTCTATAGATACACAAACACCAATTAAAACCCTG
CTTTGGAAAGTCTAACTTTTTCCCCACAG

Domain d

GCACCCGTGATCGTGATAACTACGTTGAAGAAGTTACTGGGGCCAGTCATATCAGGAAGAA
TTTGAACGACCTCAATACCGGAGAAATGGAAAGCCTTAGAGCTGCTTTCCTGCATATTCAG
GACGACGGAACATATGAATCTATTGCCCAGTACCATGGCAAACCAGGCAAATGTCAATTGA
ATGATCATAATATTGCGTGTTGTGTCCATGGTATGCCTACCTTCCCCCAGTGGCACAGACT
GTATGTGGTTCAGGTGGAGAATGCTCTCCTAAACAGGGGATCTGGTGTGGCTGTTCCTTAC

REPLACEMENT SHEET NO. 9

TGGGAGTGGACTGCTCCCATAGACCATCTACCTCATTTCATTGATGATGCAACATACTTCA
ATTCCCGACAACAGCGGTACGACCCTAACCCTTTCTTCAGGGGAAAGGTTACTTTTGAAAA
CGCAGTCACAACAAGGGACCCACAAGCCGGGCTCTTCAACTCAGATTATATGTATGAGAAT
GTTTTACTTGCACTGGAGCAGGAAAATTATTGTGACTTTGAAATTCAGTTTGAGCTTGTTC
ATAACGCACTTCATTCCATGCTGGGAGGTAAAGGGCAGTACTCCATGTCCTCCCTGGACTA
TTCTGCGTTTGATCCCGTCTTCTTCCTACATCATGCCAACACGGACAGACTGTGGCAATC
TGGCAGGAACTACAAAGATTCCGAGAACTGCCTTATGAAGAAGCGAACTGTGCAATCAACC
TCATGCATCAACCACTGAAGCCGTTCAGTGATCCACATGAGAATCACGACAATGTCACTTT
GAAATACTCAAAACCACAGGACGGATTCGACTACCAGAACCACTTCGGATACAAGTATGAC
AACCTTGAGTTCCATCACTTATCTATCCCAAGTCTTGATGCTACCCTGAAGCAAAGGAGAA
ATCACGACAGAGTGTTTGCGGGCTTCCTTCTTCATAACATAGGAACTTCTGCTGACATAAC
TATCTACATATGTCTGCCTGACGGACGGCGTGGCAATGACTGCAGTCATGAGGCGGGAACA
TTCTATATCCTCGGAGGCGAAACAGAGATGCCTTTTATCTTTGACCGTTTGTATAAATTTG
AAATCACCAAACCACTGCAACAGTTAGGAGTCAAGCTGCATGGTGGAGTTTTCGAACTGGA
GCTTGAGATCAAGGCATACAACGGTTCCTATCTGGATCCCCATACCTTTGATCCAACTATC
ATCTTTGAACCTGGAACAG

Intron d/e

GTAATGCCATCTTAATACAGTTCGTTCGTTAAATTATATATGTTCGTTTACAACACCATAC
CTTGAATTGAGGTAATACATCACTTGATATTGATAATGTAATGGTAATTGTTCTTGTTTGT
AAAACCGTTTCTGGGGTGTTTATTCACTATCCACCTGGTGGATAGTGAGTAAACACATTCG
GTTTAATATGGGTATCTAATGGACAGTGAAGTGTGCTGGCTAGGCAGATACCTTGGTTTCT
GTGAATGGAGGTAGTAGAAAGGGGTTTTGATGATTGCAG

Domain e

ATACCCATATCTTGGACCACGACCATGAGGAAGAGATACTTGTCAGGAAGAATATAATTGA
TTTGAGCCCAAGGGAGAGGGTTTCTCTAGTCAAAGCTTTGCAAAGAATGAAGAATGATCGC
TCCGCTGATGGGTACCAAGCCATTGCCTCTTTCCATGCCCTGCCACCACTCTGTCCCAATC
CATCTGCAGCTCACCGTTATGCTTGCTGTGTCCATGGCATGGCTACATTTCCCCAGTGGCA
CAGACTGTACACTGTTCAGGTTCAGGATGCCCTGAGGAGACATGGTTCACTTGTTGGTATT
CCTTACTGGGACTGGACAAAACCAGTCAACGAGTTACCCGAGCTTCTTTCTTCAGCAACAT
TTTATCATCCAATCCGGAATATTAATATTTCAAATCCATTCCTCGGGGCTGACATAGAATT
TGAAGGACCGGGCGTTCATACAGAGAGGCACATAAATACTGAGCGCCTGTTTCACAGTGGG
GATCATGACGGATACCACAACTGGTTCTTCGAAACTGTTCTCTTTGCTTTGGAACAGGAAG
ATTACTGCGATTTTGAAATACAATTTGAGATAGCCCATAATGGCATCCACACATGGATTGG
TGGAAGCGCAGTATATGGCATGGGACACCTTCACTATGCATCATATGATCCAATTTTCTAC
ATCCACCATTCACAGACGGACAGAATATGGGCTATTTGGCAAGAGCTGCAGAAGTACAGGG
GTCTATCTGGTTCGGAAGCAAACTGTGCCATTGAACATATGAGAACACCCTTGAAGCCTTT
CAGCTTTGGGCCACCCTACAATTTGAATAGTCATACGCAAGAATATTCAAAGCCTGAGGAC
ACGTTTGACTATAAGAAGTTTGGATACAGATATGATAGTCTGGAATTGGAGGGGCGATCAA
TTTCTCGCATTGATGAACTTATCCAGCAGAGACAGGAGAAAGACAGAACTTTTGCAGGGTT
CCTCCTTAAAGGTTTTGGTACATCCGCATCTGTGTCATTGCAAGTTTGCAGAGTTGATCAC
ACCTGTAAAGATGCGGGCTATTTCACTATTCTGGGAGGATCAGCCGAAATGCCATGGGCAT
TCGACAGGCTTTATAAGTATGACATTACTAAAACTCTTCACGACATGAACCTGAGGCACGA
GGACACTTTCTCTATAGACGTAACTATCACGTCTTACAATGGAACAGTACTCTCGGGAGAC
CTCATTCAGACGCCCTCCATTATATTTGTACCTGGACGCC

Intron e/f

GTGAGTACCTGTTTGCACTAAGACTTCTGTAGGCTAAAAGTGTAAGAAATATCAATTAATT
TCAATTCACCCAAACTTGAAAACGGTACCTATATAGGTTAACTTTTTGTCTACAGTAAACT
GAACATACCTACACATTTCATGAAATGATCTCTCAATATTTTCCACCAACAG

Domain f

ATAAACTCAACTCACGGAAACATACACCTAACAGAGTCCGCCATGAGCTAAGTAGCCTTAG
TTCCCGTGACATAGCAAGCTTGAAGGCAGCTTTGACAAGCCTTCAACATGATAATGGACT
GATGGTTATCAAGCTATTGCTGCCTTCCATGGCGTTCCTGCGCAGTGCCACGAGCCATCTG
GACGTGAG

Intron f(1)

GTAAATTTACAGAGCTTTATGAAGTGTGTTCAGAGTGAAGAGACCAAGATATACTTATACC
CAAAACTAGCTAGCAACAGACGATTTCACTTGTTTCGGACACTTTGTATTATACGTTGGAT
CCCAAGGTAAACGGAAACGTAACCGAGAATCAGTCCGTAAAGTGAGTGAGTGAGTTTGGGG
CTTAACGTCGCACTCAGCAATACCCCAGCTATGTGGCGACTCTCAGATTTACTGCTGGAGG
AGAACCTACATAGCCCGGTTTAACCCGTGTGGTATGTAGTAAGACCAGCGCGGCATGGCTG
GTATCTGACGGACGAAGGGTGGCGCTGCACGTATTCCAGTGGTACAACACTGCACCCCAAT
TTCACCGACCGGAGAACTGATCTCCCCTTCGGAGATATCGCCTGCCTTCCACGGGATTCGA
ACTCGGTGACCTTCAAGCCAGCGCGCTTCTAGCGGGGGCGATTAGAGGTTNAAGGCCGACG
GCTCTACCACCTTAACTATCCCCCGGCCCCACTCCTGACGGAAATGTTTATAATTCAGCCT
TTGTTTTCTTATTAAACACTCTTGGCAGATTTTCTATAGATAATGGATTCACATGTAGACA
GTCTCCCATTGTTGTAACTGGTAGTCAAGAGTTAGAATCTGAATACATTCTCCAAGATGGA
TCAAGGAAAACAATAATTACTTGATGTTGCAG

Domain f(2)

ATCGCCTGTTGCATCCACGGCATGGCGACGTTTCCTCACTGGCACCGGTTGTACACTCTGC
AGTTGGAGCAAGCGCTGCGCAGACACGGGTCCAGTGTTGCTGTTCCATACTGGGACTGGAC
CAAGCCAATCACCGAACTGCCACACATTCTGACAGACGGAGAATATTATGACGTTTGGCAA
AATGCCGTCTTGGCCAATCCGTTTGCAAGAGGTTATGTGAAAATTAAAGATGCATTTACGG
TGAGAAATGTCCAGGAAAGTCTGTTCAAAATGTCAAGTTTTGGAAAGCACTCGCTTCTGTT
TGACCAGGCTTTGTTGGCTCTTGAACAAACTGACTACTGTGACTTCGAAGTTCAGTTTGAA
GTGATGCATAACACGATCCATTATCTCGTAGGAGGGCGTCAAACGTACGCCTTCTCCTCTC
TCGAGTATTCCTCATACGATCCAATCTTCTTTATTCACCACTCGTTTGTTGACAAAATATG
GGCTGTATGGCAAGAACTGCAAAGCAGGAGACATCTACAGTTTAGAACAGCTGATTGTGCT
GTGGGCCTCATGGGTCAGGCAATGAGGCCTTTCAACAAGGATTTCAACCACAACTCGTTCA
CCAAGAAGCACGCAGTCCCTAATACAGTATTTGATTATGAAGATCTTGGCTATAACTATGA
CAACCTTGAAATCAGTGGTTTAAACTTAAATGAGATCGAGGCGTTAATAGCAAAACGCAAG
TCACATGCTAGAGTCTTTGCTGGGTTCCTGTTGTTTGGATTAGGAACTTCGGCTGATATAC
ATCTGGAAATTTGCAAGACATCGGAAAACTGCCATGATGCTGGTGTGATTTTCATCCTTGG
AGGTTCTGCAGAGATGCATTGGGCATACAACCGCCTCTACAAGTATGACATTACAGAAGCA
TTGCAGGAATTTGACATCAACCCTGAAGATGTTTTCCATGCTGATGAACCATTTTTCCTGA
GGCTGTCGGTTGTTGCTGTGAATGGAACTGTCATTCCATCGTCTCATCTTCACCAGCCAAC
GATAATCTATGAACCAGGCGAAG

Intron f/g

GTGAGATATATGCAAATTGAATGTTGTCCAGATGCGTTGTTTACATTTATATGCTTGGAAT
TGTCCTGAACGAATACAGTGGAATAACCAAAAGCTGAAAAATAAAAAGATATATACTTCAT
TCTGAATTTGTCAGTATTGCTGACCCAAAAACACGTTATCCATGTCGACACTATATTTGCC
TTTCTGAATCTGAGACTGCGTTATGTTTCTAATAATCACGAAATATGGTATACAGGTTGTG
TATCTGTAGAATACCCAAGGCAGAATTTAAAGGTCACACCCTGTTTAATACAG

Domain g(1)

ATCACCATGACGACCATCAGTCGGGAAGCATAGCAGGATCCGGGGTCCGCAAGGACGTGAA
CACCTTGACTAAGGCTGAGACCGACAACCTGAGGGAGGCGCTGTGGGGTGTCATGGCAGAC
CACGGTCCCAATGGCTTTCAAGCTATTGCTGCTTTCCATGGAAAACCAGCTTTGTGTCCCA
TGCCTGATGGCCACAACTACTCATGTTGTACTCACGGCATGGCTACCTTCCCACACTGGCA
TCGCCTCTACACCAAGCAGATGGAGGATGCAATGAGGGCGCATGGGTCTCATGTCGGCCTG
CCCTACTGGGACTGGACTGCTGCCTTCACCCACCTGCCAACACTGGTCACCGACACGGACA
ACAACCCCTTCCAACAT

Intron g(2)

GTAAGAGCGGGGTAGGGATGGGGTGGTAGGGGGTGGGTTGTTCTATTACTTCCCGCTTCAC
TTGTATGAAATGGATAACCTTGGCTGCATCCCAATTGCGTGATCGATTCTCTTTCGATTCA
CTCGTGCGATTAGACTGCCTTATTTACTATAGTAGTTAGAATGTTGCTCAGTGCGCCGTTA
AACAACTAATACACAAAACCGCATTTGTTTTATATGGTCACTCTACTGTTTATCACGTATA
TGTATGTTCCGACTCACTGGTTGGTGCGTACCATTCTACTGTCACACTGAGAGCCAATGTT
CTCAGATGTGTGAAATGTTTGAAAGCCGTTTCTACATAATATTGCAGGAATACCATTGTAG
AATGTAGTCAAACAGGTAACAATCTGTTAGTGAGCCCAGTTCGAGGTTGCGTTGTAGGGTG
TAGTCCAACAGGTAGGCAGTCCATAAGCATAGTTTTTAAGCATTTTAGATCATCTATAATT
AACCACATGGTTAGCCGCTATGTTTAGTTTAATCCAGTATAAGTTAGAACTGTTATATTTC
GAAGGGAAGTGAGTAAATCCTTATTCCTTGACTACCATTTAATAGATTTCCCAATGACTCC
ATTCAACTCCTAACTTTCACATCACTGCTCTCTTCAACAG

Domain g(2)

GGACACATTGATTATCTCAATGTCAGCACAACTCGATCTCCCCGAGACATGCTGTTCAACG
ACCCCGAGCATGGATCAGAGTCGTTCTTCTACAGACAAGTCCTCTTAGCTCTGGAACAAAC
TGATTTCTGCAAATTCGAAGTTCAGTTTGAGATAACCCACAATGCCATCCATTCCTGGACA
GGTGGCCACAGCCCCTACGGAATGTCCACTCTCGACTTCACTGCCTACGATCCTCTCTTCT
GGCTTCACCACTCCAACACCGACAGAATCTGGGCTGTCTGGCAAGCTTTGCAAGAATACAG
AGGACTTCCATACAACCATGCCAATTGTGAGATCCAGGCAATGAAAACGCCCCTGAGGCCT
TTCAGTGACGATATCAACCACAACCCAGTCACAAAGGCTAACGCGAAGCCATTAGATGTGT
TCGAGTATAATCGGTTGAGCTTCCAGTACGACAACCTCATCTTCCATGGATACAGTATTCC
GGAACTTGATCGCGTGCTTGAAGAAAGAAAGGAGGAGGACAGAATATTTGCTGCCTTCCTT
CTCAGTGGAATCAAGCGTAGTGCTGATGTAGTGTTCGACATATGCCAGCCAGAACACGAAT
GTGTGTTCGCAGGGACTTTTGCGATTTTGGGAGGGGAGCTAGAAATGCCCTGGTCCTTCGA
CAGACTGTTCCGCTATGATATCACCAAGGTGATGAAGCAGCTACACCTGAGGCATGACTCT
GACTTTACCTTCAGGGTGAAGATTGTCGGCACCGACGACCACGAGCTTCCTTCAGACAGTG
TCAAAGCACCAACTATTGAATTTGAACCGGGCG

Intron g/h

GTGAGTACGACAGGCATTTCTAGTAAAAACCTACTTTTGGTAAAAGGTTCGAGAAATCACT
TGAAGCAACAACATGATTTTGTAACGCCTATTACACGTGAACATGTCACACCCGGTGATGC
CGTTTAATGGACATGCCTCTGTTAATGAAAGGGGTAAGTACATGTGTATGGGGATGGGATG
GGAGCCACCTGTCCCAATTTCATAGGTCCCTAGGATCCCAGTTGCGTAGGAATCCCCTGAT
TAATGCCTTGTGAATTCCTCCTGGAATTGTCCTGGCCCAAATTTTTACAAACCCGCCCCGA
TATACCTTGGAAATAATTGGGCCTAAGGGTGGGGCTTTTAAGGACCAAGAACCCAACCTAA
ACCCCAACCCATTTTTTCCCACCCATTCCAGGTTTTGTTTTACCAAATAAAAAGGTTTCCA
CTTTGAGGAAACCCTTTAAGGGTTCTTTTCAGGGCTTTTTTTCTTTTCTGGAATTCCAAT
TCCGGGGGAACAAAATACATATATTTCACAGACCTTTGGTCAAATTTATATAATTTCCGAC
TTCATGTCATAGGTTTGTCTTTCTTCCTACACAG

Domain h

TGCACAGAGGCGGAAACCACGAAGATGAACACCATGATGACAGACTCGCAGATGTCCTGAT
CAGGAAAGAAGTTGACTTCCTCTCCCTGCAAGAGGCCAACGCAATTAAGGATGCACTGTAC
AAGCTCCAGAATGACGACAGTAAAGGGGGCTTTGAGGCCATAGCTGGCTATCACGGGTATC
CTAATATGTGTCCAGAAAGAGGTACCGACAAGTATCCCTGCTGTGTCCACGGAATGCCCGT
GTTCCCCCACTGGCACCGCCTGCATACCATTCAGATGGAGAGAGCTCTGAAAAACCATGGC
TCTCCAATGGGCATTCCTTACTGGGATTGGACAAAGAAGATGTCGAGTCTTCCATCTTTCT
TTGGAGATTCCAGCAACAACAACCCTTTCTACAAATATTACATCCGGGGCGTGCAGCACGA
AACAACCAGGGACATTAATCAGAGACTCTTTAATCAAACCAAGTTTGGTGAATTTGATTAC
CTATATTACCTAACTCTGCAAGTCCTGGAGGAAAACTCGTACTGTGACTTTGAAGTTCAGT
ATGAGATCCTCCATAACGCCGTCCACTCCTGGCTTGGAGGAACTGGAAAGTATTCCATGTC
TACCCTGGAGCATTCGGCCTTTGACCCTGTCTTCATGATTCACCACTCGAGTTTGGATAGA
ATCTGGATCCTTTGGCAGAAGTTGCAAAAGATAAGAATGAAGCCTTACTACGCATTGGATT
GTGCTGGCGACAGACTTATGAAAGACCCCCTGCATCCCTTCAACTACGAAACCGTTAATGA
AGATGAATTCACCCGCATCAACTCTTTCCCAAGCATACTGTTTGACCACTACAGGTTCAAC
TATGAATACGATAACATGAGAATCAGGGGTCAGGACATACATGAACTTGAAGAGGTAATTC
AGGAATTAAGAAACAAAGATCGCATATTTGCTGGTTTTGTTTTGTCGGGCTTACGGATATC
AGCTACAGTGAAAGTATTCATTCATTCGAAAAACGATACAAGTCACGAAGAATATGCAGGA
GAATTTGCAGTTTTGGGAGGTGAGAAGGAGATGCCGTGGGCATATGAAAGAATGCTGAAAT
TGGACATCTCCGATGCTGTACACAAGCTTCACGTGAAAGATGAAGACATCCGTTTTAGAGT
GGTTGTTACTGCCTACAACGGTGACGTTGTTACCACCAGGCTGTCTCAGCCATTCATCGTC
CACCGTCCAGCCCATGTGGCTCACGACATCTTGGTAATCCCAGTAGGTGCGGGCCATGACC
TTCCGCCTAAAGTCGTAGTAAAGAGCGGCACCAAAGTCGAGTTTACACCAATAGATTCGTC
GGTGAACAAAGCAATGGTGGAGCTGGGCAGCTATACTGCTATGGCTAAATGCATCGTTCCC
CCTTTCTCTTACCACGGCTTTGAACTGGACAAAGTCTACAGCGTCGATCACGGAGACTACT
ACATTGCTGCAGGTACCCACGCGTTGTGTGAGCAGAACCTCAGGCTCCACATCCACGTGGA
ACACGAGTAG

3´UTR

TTCACAG

Intron UTR

GTGAGGAGAAGGCCCCAGGCTAGCAGGGCAATGGATGAAGGAAATAGGGGCAAAGGGAATA
GCAGTTACACCATCGACATTTCCAACCTCCTCAGAAACTAATATATAGCCTTAATACAACC
AGCCAAGACTCAACGGGCAGCCGGGGTGGGGGATTTGGTGGTCGCTGTTTCAGACCAGGG
TGCAAAATATCAGTGCGCAAATCAACATGTTGCGTGTCAGACACTGACACAGCAGTCATTG
AACCTGCAGACCCATAACAGGAAAATGGGGCAGATACGATCAAAGACAGTGTAAAATAGGG
ATAAGTAGGCATATGCAACCACCTGATGGAAATGAAAAGGGGTAAGTTTAAACCCCGGCTA
CCAAAGGTCCAATGGTTCCTTAACCCAGCTTACGCTATCCCTCTAATTTCAGTATTGAGCT
GATTTCTGTCGAGTTCATGTAAACTGTATACTTTCTGTATTATTACAG

3´UTR

GTTGCTATGCCGACTGCGCTATATTGGTGAACGAGACGATGAGGACATCTCTGAAAGAGTT
CGCCAAGTGATGTGTAGGTCACGGAAGTATTGTTGAGCTAACAATATGATGATTTCAAAAT
GACTTGGCGCTCTAGGACAAAGACATAATTCATCAGCACCCTGTGCACCAACTCTTTGTTT
GCTGCAAACGTCTGACAAGCGACACGTCAATCAACAAGCTGTTCAAACTCAAGTGGATGTA
ACTAGAATCGTTGGGCCATCGTTCACAAAGTATTGACAGATGTCACACATGATGGCGAGAA
ACACTTTAGAACTTTTAATGACCTAGAGTGACTTGTAAATATGTAAATATATTCTTCAAAG
ACTCAGCTGAACTATTGTTGGATAACACATCAATTCCCTCAACAAAATGCTTTATCTTCAC
ATGGATGTATGTAATGTGGCCGGCAATAAAGTATATATATGTATAAAAAAAAAAAAAAAA
A

Figure 5

Derived primary structure of HtH1

Signal peptide

LVQFLLVALVAGAGA

Domain a

DNVVRKDVSHLTDDEVQALHGALHDVTASTGPLSFEDITSYHAAPASCDYKGRKIACCVHG
MPSFPFWHRAYVVQAERALLSKRKTVGMPYWDWTQTLTHLPSLVTEPIYIDSKGGKAQTNY
WYRGEIAFINKKTARAVDDRLFEKVEPGHYTHLMETVLDALEQDEFCKFEIQFELAHNAIH
YLVGGKFEYSMSNLEYTSYDPIFFLHHSNVDRLFAIWQRLQELRGKNPNAMDCAHELAHQQ
LQPFNRDSNPVQLTKDHSTPADLFDYKQLGYSYDSLNLNGMTPEQLKTELDERHSKERAFA
SFRLSGFGGSANVVVYACVPDDDPRSDDYCEKAGDFFILGGQSEMPWRFYRPFFYDVTEAV
HHLGVPLSGHYYVKTELFSVNGTALSPDLLPQPTVAYRPGK

Domain b

GHLDPPVHHRHDDDLIVRKNIDHLTREEEYELRMALERFQADTSVDGYQATVEYHGLPARC
PRPDAKVRFACCMHGMASFPHWHRLFVTQVEDALVRRGSPIGVPYWDWTKPMTHLPDLASN
ETYVDPYGHTHHNPFFNANISFEEGHHHTSRMIDSKLFAPVAFGEHSHLFDGILYAFEQED
FCDFEIQFELVHNSIHAWIGGSEDYSMATLHYTAFDPIFYLHHSNVDRLWAIWQALQIRRH
KPYQAHCAQSVEQLPMKPFAFPSPLNNNEKTHSHSVPTDIYDYEEVLHYSYDDLTFGGMNL
EEIEEAIHLRQQHERVFAGFLLAGIGTSALVDIFINKPGNQPLKAGDIAILGGAKEMPWAF
DRLYKVEITDSLKTLSLDVDGDYEVTFKIHDMHGNALDTDLIPHAAVVSEPAH

Domain c

PTFEDEKHSLRIRKNVDSLTPEETNELRKALELLENDHTAGGFNQLGAFHGEPKWCPNPEA
EHKVACCVHGMAVFPHWHRLLALQAENALRKHGYSGALPYWDWTRPLSQLPDLVSHEQYTD
PSDHHVKHNPWFNGHIDTVNQDTTRSVREDLYQQPEFGHFTDIAQQVLLALEQDDFCSFEV
QYEISHNFIHALVGGTDAYGMASLRYTAYDPIFFLHHSNTDRIWAIWQSLQKYRGKPYNTA
NCAIESMRRPLQPFGLSSAINPDRITREHAIPFDVFNYRDNLHYVYDTLEFNGLSISQLDR
ELEKIKSHERVFAGFLLSGIKKSALVKFEVCTPPDNCHKAGEFYLLGDENEMAWAYDRLFK
YDITQVLEANHLHFYDHLFIRYEVFDLKGVSLGTDLFHTANVVHDSGT

Domain d

GTRDRDNYVEEVTGASHIRKNLNDLNTGEMESLRAAFLHIQDDGTYESIAQYHGKPGKCQL
NDHNIACCVHGMPTFPQWHRLYVVQVENALLNRGSGVAVPYWEWTAPIDHLPHFIDDATYF
NSRQQRYDPNPFFRGKVTFENAVTTRDPQAGLFNSDYMYENVLLALEQENYCDFEIQFELV
HNALHSMLGGKGQYSMSSLDYSAFDPVFFLHHANTDRLWAIWQELQRFRELPYEEANCAIN
LMHQPLKPFSDPHENHDNVTLKYSKPQDGFDYQNHFGYKYDNLEFHHLSIPSLDATLKQRR
NHDRVFAGFLLHNIGTSADITIYICLPDGRRGNDCSHEAGTFYILGGETEMPFIFDRLYKF
EITKPLQQLGVKLHGGVFELELEIKAYNGSYLDPHTFDPTIIFEPGT

Domain e

DTHILDHDHEEEILVRKNIIDLSPRERVSLVKALQRMKNDRSADGYQAIASFHALPPLCPN
PSAAHRYACCVHGMATFPQWHRLYTVQVQDALRRHGSLVGIPYWDWTKPVNELPELLSSAT
FYHPIRNINISNPFLGADIEFEGPGVHTERHINTERLFHSGDHDGYHNWFFETVLFALEQE
DYCDFEIQFEIAHNGIHTWIGGSAVYGMGHLHYASYDPIFYIHHSQTDRIWAIWQELQKYR
GLSGSEANCAIEHMRTPLKPFSFGPPYNLNSHTQEYSKPEDTFDYKKFGYRYDSLELEGRS
ISRIDELIQQRQEKDRTFAGFLLKGFGTSASVSLQVCRVDHTCKDAGYFTILGGSAEMPWA
FDRLYKYDITKTLHDMNLRHEDTFSIDVTITSYNGTVLSGDLIQTPSIIFVPGR

Domain f

HKLNSRKHTPNRVRHELSSLSSRDIASLKAALTSLQHDNGTDGYQAIAAFHGVPAQCHEPS
GREIACCIHGMATFPHWHRLYTLQLEQALRRHGSSVAVPYWDWTKPITELPHILTDGEYYD
VWQNAVLANPFARGYVKIKDAFTVRNVQESLFKMSSFGKHSLLFDQALLALEQTDYCDFEV
QFEVMHNTIHYLVGGRQTYAFSSLEYSSYDPIFFIHHSFVDKIWAVWQELQSRRHLQFRTA
DCAVGLMGQAMRPFNKDFNHNSFTKKHAVPNTVFDYEDLGYNYDNLEISGLNLNEIEALIA
KRKSHARVFAGFLLFGLGTSADIHLEICKTSENCHDAGVIFILGGSAEMHWAYNRLYKYDI
TEALQEFDINPEDVFHADEPFFLRLSVVAVNGTVIPSSHLHQPTIIYEPGE

Domain g

DHHDDHQSGSIAGSGVRKDVNTLTKAETDNLREALWGVMADHGPNGFQAIAAFHGKPALCP
MPDGHNYSCCTHGMATFPHWHRLYTKQMEDAMRAHGSHVGLPYWDWTAAFTHLPTLVTDTD
NNPFQHGHIDYLNVSTTRSPRDMLFNDPEHGSESFFYRQVLLALEQTDFCKFEVQFEITHN
AIHSWTGGHSPYGMSTLDFTAYDPLFWLHHSNTDRIWAVWQALQEYRGLPYNHANCEIQAM
KTPLRPFSDDINHNPVTKANAKPLDVFEYNRLSFQYDNLIFHGYSIPELDRVLEERKEEDR
IFAAFLLSGIKRSADVVFDICQPEHECVFAGTFAILGGELEMPWSFDRLFRYDITKVMKQL
HLRHDSDFTFRVKIVGTDDHELPSDSVKAPTIEFEPG

Domain h

VHRGGNHEDEHHDDRLADVLIRKEVDFLSLQEANAIKDALYKLQNDDSKGGFEAIAGYHGY
PNMCPERGTDKYPCCVHGMPVFPHWHRLHTIQMERALKNHGSPMGIPYWDWTKKMSSLPSF
FGDSSNNNPFYKYYIRGVQHETTRDVNQRLFNQTKFGEFDYLYYLTLQVLEENSYCDFEVQ
YEILHNAVHSWLGGTGQYSMSTLEYSAFDPVFMIHHSSLDRIWILWQKLQKIRMKPYYALD
CAGDRLMKDPLHPFNYETVNEDEFTRINSFPSILFDHYRFNYEYDNMRIRGQDIHELEEVI
QELRNKDRIFAGFVLSGLRISATVKVFIHSKNDTSHEEYAGEFAVLGGEKEMPWAYERMLK
LDISDAVHKLHVKDEDIRFRVVVTAYNGDVVTTRLSQPFIVHRPAHVAHDILVIPVGAGHD
LPPKVVVKSGTKVEFTPIDSSVNKAMVELGSYTAMAKCIVPPFSYHGFELDKVYSVDHGDY
YIAAGTHALCEQNLRLHIHVEHE

Figure 6

HtH2 cDNA sequence and intron structure

Domain b

CACAGACTGTTCGTCACCCAGGTGGAAGATGCTCTGATCAGGCGAGGATCGCCTATAGGGG
TCCCCTACTGGGACTGGACTCAGCCTATGGCGCATCTCCCAGGACTTGCAGACAACGCCAC
CTATAGAGATCCCATCAGCGGGGACAGCAGACACAACCCCTTCCACGATGTTGAAGTTGCC
TTTGAAAATGGACGTACAGAACGTCACCCAGATAGTAGATTGTTTGAACAACCTTTATTTG
GCAAACATACGCGTCTCTTCGACAGTATAGTCTATGCTTTTGAGCAGGAGGACTTCTGCGA
TTTTGAAGTTCAATTTGAGATGACCCATAATAATATTCACGCCTGGATTGGTGGCGGCGAG
AAGTATTCCATGTCTTCTCTACACTACACAGCCTTCGACCCTATCTTCTACCTTCGTCACT
CCAACACTGACCGGCTCTGGGCAATTTGGCAAGCGTTGCAGATACGAAGAAACAGGCCTTA
CAAGGCTCATTGTGCTTGGTCTGAGGAACGCCAGCCTCTCAAACCTTTCGCCTTCAGTTCC
CCACTGAACAACAACGAAAAAACCTACGAAAACTCGGTGCCCACCAACGTTTACGACTACG
AAGGAGTCCTTGGCTATACTTATGATGACCTCAACTTCGGGGGCATGGACCTGGGTCAGCT
TGAGGAATACATCCAGAGGCAGAGACAGAGAGACAGGACCTTTGCTGGTTTCTTTCTGTCA
CATATTGGTACATCAGCGAATGTTGAAATCATTATAGACCATGGGACTCTTCATACCTCCG
TGGGCACGTTTGCTGTTCTTGGCGGAGAGAAGGAGATGAAATGGGGATTTGACCGTTTGTA
CAAATATGAGATTACAGATGAACTGAGGCAACTTAATCTCCGTGCTGATGATGTTTTCAGC
ATCTCTGTTAAAGTAACTGATGTTGATGGCAGTGAGCTGTCCTCTGAACTCATCCCATCTG
CTGCTATCATCTTCGAACGAAGCCATA

Intron b/c

GTAAGTAGCTACCTGTTTATTCAATTTTTTCGCTTTGCCAATCAATTCATTCAGCTTGAAA
TTCAATAATTGTGTTTTGCATGGCTGAAAACCAATTTGAACTCTTTTCTTTTCTCAGGTCG
AACTCAAATAAATAATCACTAATTGTTATGCACGCGGGTAGGGCATACATACTATATCCAC
ATCGGTCATCTCAAAATGCAAACAAATTGTCTTATTTCCGTTGGGACAAGCAAACCCCCTT
TCCTGTAATCTTGCCTTTGGCATCCACTGGAATTAATGTTGACTGGTAATTGATACTGGCT
CTCTTCTTGCATAGAGTTAATATCTATAGTTTGTAAATCTTTATGATTTTGCTATTTATAT
TTCGACAGCATGCTATAGACACCCTAGACTATTGTATAGCCACTTGTATTGTTTTTCCATT
TATTATTTATAACAGAACATGGCTTGTAATTTTTATTTACCTTCCAG

Domain c

TTGACCATCAGGACCCGCATCATGACACAATCATTAGGAAAAATGTTGATAATCTTACACC
CGAGGAAATTAATTCTCTGAGGCGGGCAATGGCAGACCTTCAATCAGACAAAACCGCCGGT
GGATTCCAGCAAATTGCTGCTTTTCACGGGGAACCCAAATGGTGCCCAAGTCCCGATGCTG
AGAAGAAGTTCTCCTGCTGTGTCCATGGAATGGCTGTCTTCCCTCACTGGCACAGACTCCT
GACCGTGCAAGGCGAGAATGCCCTGAGAAAGCATGGATGTCTCGGAGCTCTCCCCTACTGG
GACTGGACTCGGCCCCTGTCTCACCTACCTGATTTGGTTTTGGTAAGTAGCAGAACTACAC
CGATGCCATATTCCACCGTGGAAGCCCGAAACCCCTGGTACAGCGGCCATATTGATACAGT
TGGTGTTGACACAACAAGAAGCGTCCGTCAAGAACTGTATGAAGCTCCTGGATTTGGCCAT
TATACTGGGGTCGCTAAGCAAGTGCTTCTGGCTTTGGAGCAGGATGACTTCTGTGATTTTG
AAGTCCAGTTTGAGATAGCTCACAATTTCATTCACGCTCTTGTCGGCGGAAGCGAGCCATA
TGGTATGGCGTCACTCCGTTACACTACTTATGATCCAATTTTCTACCTCCATCATTCTAAC
ACTGACAGACTCTGGGCTATATGGCAGGCTCTACAAAAGTACAGGGGCAAACCTTACAATT
CCGCCAACTGCGCCATTGCTTCTATGAGAAAACCCCTACAACCCTTTGGTCTGACTGATGA
GATCAACCCGGATGATGAGACAAGACAGCATGCTGTTCCTTTCAGTGTCTTTGATTACAAG
AACAACTTCAATTATGAATATGACACCCTTGACTTCAACGGACTATCAATCTCCCAGCTGG
ACCGTGAACTGTCACGGAGAAAGTCTCATGACAGAGTATTTGCCGGATTTTTGCTGCATGG

TATTCAGCAGTCTGCACTAGTTAAATTCTTTGTCTGCAAATCAGATGATGACTGTGACCAC
TATGCTGGTGAATTCTACATCCTTGGTGATGAAGCTGAAATGCCATGGGGCTATGATCGTC
TTTACAAATATGAGATCACTGAGCAGCTCAATGCCCTGGATCTACACATCGGAGATAGATT
CTTCATCAGATACGAAGCGTTTGATCTTCATGGTACAAGTCTTGGAAGCAACATCTTCCCC
AAACCTTCTGTCATACATGACGAAGGGGCAG

Intron c/d

GTGAGAACATTGATAATAGTTCAAATgAAGTATATCCGATTCAAGCTGTCGATACAAGATg
AGATACATAATCACAATGTTTGTATTAGATATCTCTCTTAATTTAATGCCGCTTTTATCAA
TATTCGAGCAATCCTTCAGCAACATACACCAGCAAATGTTTCATCAACAGACTATATTATT
TAATCTTTTAAAAATCCTTTTCTGTTGTTATAAATACTTAAAGTATCGAATTCCTTGAATG
CGTCTTCTCTGCAGCATATAGTTAAGTTGTTGTGTTTCTCTGTCAG

Domain d

GTCACCATCAGGCTGACGAGTACGACGAAGTTGTAACTGCTGCAAGCCACATCAGAAAGAA
TTTAAAAGATCTGTCAAAGGGAGAAGTAGAGAGCCTAAGGTCTGCCTTCCTGCAACTTCAG
AACGACGGAGTCTATGAGAATATTGCCAAGTTCCACGGCAAGCCTGGGTTGTGTGATGATA
ACGGTCGCAAGGTTGCCTGTTGTGTCCATGGAATGCCCACCTTCCCCCAGTGGCACAGGCT
CTATGTCCTCCAGGTGGAGAATGCTTTGCTGGAGAGAGGATCTGCCGTCTCTGTGCCATAC
TGGGACTGGACTGAAACATTTACAGAGCTGCCATCTTTGATTGCTGAGGCTACCTATTTCA
ATTCCCGTCAACAAACGTTTGACCCTAATCCTTTCTTCAGAGGTAAAATCAGTTTTGAGAA
TGCTGTTACAACACGTGATCCCCAGCCTGAGCTGTACGTTAACAGGTACTACTACCAAAAC
GTCATGTTGGTTTTTGAACAGGACAACTACTGCGACTTCGAGATACAGTTTGAGATGGTTC
ACAATGTTCTCCATGCTTGGCTTGGTGGAAGAGCTACTTATTCTATTTCTTCTCTTGATTA
TTCTGCATTCGACCCTGTGTTTTTCCTTCACCATGCGAACACAGATAGATTGTGGCCATC
TGGCAGGAGCTGCAGAGGTACAGGAAGAAGCCATACAATGAAGCGGATTGTGCCATTAACC
TAATGCGCAAACCTCTACATCCCTTCGACAACAGTGATCTCAATCATGATCCTGTAACCTT
TAAATACTCAAAACCCACTGATGGCTTTGACTACCAGAACAACTTTGGATACAAGTATGAC
AACCTTGAGTTCAATCATTTCAGTATTCCCAGGCTTGAAgAAATCATTCGtATTAGACAAC
GTCAAGATCGTGTGTTTGCAGGATTCCTCCTTCACAACATTGGGACATCCGCAACTGTTGA
GATATTCGTCTGTGTCCCTACCACCAGCGGTGAGCAAAACTGTGAAAACAAAGCCGGAACA
TTTGCCGTACTCGGAGGAGAAACAGAGATGGCGTTTCATTTTGACAGACTCTACAGGTTTG
ACATCAGTGAAACACTGAGGGACCTCGGCATACAGCTGGACAGCCATGACTTTGACCTCAG
CATCAAGATTCAAGGAGTAAATGGATCCTACCTTGATCCACACATCCTGCCAGAGCCATCC
TTGATTTTTGTGCCTGGTTCAAGT

Intron d/e

AAGAAAGTTTCACTGTCTAAATCTTTTTTTATGATAGAGGGTAGAGAAGTGGAGACAATGT
GACAATATATTGAATAAAGTTGTTTAAAATTTATAACTCTCATAAGTTCATATTATGCTGA
AGCTGTAGCCATCTATAACTGTGTAACATGAAATGTTAAGACATTAACCTAAATACTTCAG
CTGATAACAAAACAATGTTAATACATACGTCAATGTAACATTTTCTTATCTTTAGGTTATA
GCATAAACACTTCAGAGATACAGTGACGAAAACCTCTATTTAAATATTTCAGGT

Domain e

TCTTTCCTGCGTCCTGATGGGCATTCAGATGACATCCTTGTGAGAAAAGAAGTGAACAGCC
TGACAACCAGGGAGACTGCATCTCTGATCCATGCTCTGAAAAGTATGCAGGAAGACCATTC
ACCTGACGGGTTCCAAGCCATTGCCTCTTTCCATGCTCTGCCACCACTCTGCCCTTCACCA
TCTGCAGCTCACCGTTATGCTTGCTGTGTCCACGGCATGGCTACATTTCCCCAGTGGCACA
GATTGTACACTGTACAGTTCCAGGATGCACTGAGGAGACATGGAGCTACGGTAGGTGTACC
GTATTGGGATTGGCTGCGACCGCAGTCTCACCTACCAGAGCTTGTCACCATGGAGACATAC

CATGATATTTGGAGTAACAGAGATTTCCCCAATCCTTTCTACCAAGCCAATATTGAGTTTG
AAGGAGAAAACATTACAACAGAGAGAGAAGTCATTGCAGACAAACTTTTTGTCAAAGGTGG
ACACGTTTTTGATAAACTGGTTCTTCAAACAAGCCATCCTAGCGCTGAGCAGGAAAACTAC
TGTGACTTTGAGATTCAGTTTGAAATTCTTCACAACGGCGTTCACACGTGGGTCGGAGGCA
GTCGTACCTACTCTATCGGACATCTTCATTACGCATTCTACGACCCTCTTTTCTACCTTCA
CCATTTCCAGACAGACCGTATTTGgGCAATCTGGCAAGAACTCCAGGAACAGAGAGGGCTC
TCGGGTGATGAGGCTCACTGTGCTCTCGAGCAAATGAGAGAACCATTGAAGCCTTTCAGCT
TCGGCGCTCCTTATAACTGGAATCAGCTCACACAGGATTTCTCCCGACCCGAGGACACCTT
CGACTACAGGAAGTTTGGTTATGAATATGACAATTTAGAATTCCTGGGAATGTCAGTTGCT
GAACTGGATCAATACATTATTGAACATCAAGAAAATGATAGAGTATTCGCTGGGTTCCTGT
TGAGTGGATTCGGAGGTTCCGCATCAGTTAATTTCCAGGTTTGTAGAGCTGATTCCACATG
TCAGGATGCTGGGTACTTCACCGTTCTTGGTGGCAGTGCTGAGATGGCGTGGGCATTTGAC
AGGCTTTACAAATATGACATTACTGAAACTCTGGAGAAAATGCACCTTCGATATGATGATG
ACTTCACAATCTCTGTCAGTCTGACCGCCAACAACGGAACTGTCCTGAGCAGCAGTCTAAT
CCCAACACCGAGTGTCATATTCCAGCGGGGACATC

Intron e/f

AAGTAGTAAACTGCTCAGATTGTTTTCATAATTACTCCACTATTAAGTAAAAAGTACTAGT
AATTCAATAGTACTGTTCACAGAGAAATGTAACACAATAGACCACAGAGTCCATTTGTTAA
ACGCCTTTGGCTTGGTAAGTCTGAGGTTTTGGTGACTGATGGAAAGCTAAAATATATTTTG
ACAG

Domain f(1)

GTGACATAAATACCAGGAGCATGTCACCGAACCGTGTTCGCCGTGAGCTGAGCGATCTGTC
TGCGAGGGACCTGTCTAGTCTCAAGTCTGCTCTGCGAGACCTACAGGAGGATGATGGCCCC
AACGGATACCAGGCTCTTGCAGCCTTCCATGGGCTACCAGCAGGCTGCCATGATAGCCGGG
GAAATGAGAT

Intron f

ATATTTAAAGTATTTTATCTTACGCATGACCCTGACCCTATTATTTTTTTAATCCTATGAT
GAAACATTTACTTAGACTGGCTTGTGAGCCCCAGGCAAAATGCACTGTAAAAATACACTGA
CAGAGGATTAGGCATTCTTGGGAGTACTGTATAGTTAGTTGCATACATATTAGCGTTCCCT
CACTAAAACGAATCTCTGAATGCTATCAATTAAAGATCATGATGCTTTGATTGTGTCTACT
GTATTTAAAATGGTGTTAAGATTTGCAATTACAATATACACAAACACGTTTCCTGCATCTC
GGAGAATGCAATCTTTCGTTGTACGCGTCTGTTTTCATATTTTTATGCATGTAGTTTGCAC
TACTTAGCGTCCAATAAATCCATTCACAAAATCACACAAACAAACGATTTTAGGAATGTGA
CTGTAGCTGCAACGAATATACCTGATCCTTTCTTGTTCCAGAT

Domain f(2)

CGCATGTTGCATTCACGGGATGCCGACCTTCCCCCAGTGGCACAGACTGTACACCCTGCAG
TTGGAGATGGCTCTGAGGAGACATGGATCATCTGTCGCCATCCCCTACTGGGACTGGACAA
AGCCTATCTCCGAACTCCCCTCGCTCTTCACCAGCCCTGAGTATTATGACCCATGGCATGA
TGCTGTGGTAAACAACCCATTCTCCAAAGGTTTTGTCAAATTTGCAAATACCTACACAGTA
AGAGACCCACAGGAGATGCTGTTCCAGCTTTGTGAACATGGAGAGTCAATCCTCTATGAGC
AAACTCTTCTTGCTCTTGAGCAAACCGACTACTGTGATTTTGAGGTACAGTTTGAGGTCCT
CCATAACGTGATCCACTACCTTGTTGGTGGACGTCAGACCTACGCATTGTCTTCTCTGCAT
TATGCCTCCTACGACCCATTCTTCTTTATACACCATTCCTTTGTGGATAAGATGTGGGTAG
TATGGCAAGCTCTTCAAAAGAGGAGGAAACTTCCATACAAGCGAGCTGACTGTgcTGTCAA
CCTAATGACTAAACCAATGAGGCCATTTGACTCCGATATGAATCAGAACCCATTCACAAAG
ATGCACGCAGTTCCCAACACACTCTATGACTACGAGACACTGTACTACAGCTACGATAATC
TCGAAATAGGTGGCAGGAATCTCGACCAGCTTCAGGCTGAAATTGACAGAAGCAGAAGCCA

```
CGATCGCGTTTTTGCTGGATTCTTGCTTCGTGGAATCGGAACTTCTGCTGATGTCAGGTTT
TGGATTTGTAGAAATGAAAATGACTGCCACAGGGGTGGAATAATTTTCATCTTAGGTGGAG
CCAAGGAAATGCCATGGTCATTTGACAGAAACTTCAAGTTTGATATCACCCATGTACTCGA
GAATGCTGGCATTAGCCCAGAGGACGTGTTTGATGCTGAGGAGCCATTTTATATCAAGGTT
GAGATCCATGCTGTTAACAAGACCATGATACCGTCGTCTGTGATCCCAGCCCCAACTATCA
TCTATTCTCCTGGGGAAG
```

Intron f/g

```
GTGAGAGAACCAGTAATAGCTACTGTCTACAAAGAATGTGTTCATTTAAAGACCTGACTGT
AGGCCGATGGCTGCTGTCATCTCCTCCGCCTCCTCCTCCTGTTCCTCCTCCGAAGGGGTCA
GCTTCAGGTTCTCTTGCCAATATGCCAAGCAGACCTCCTGAGCAGGCAGTATATATACGTA
AGGGAAGCAAGTATGGACCATCGCGCGGCATGTAGAGATACAATGATCAGCTGTCTGCTGT
TCCACTCCTGTCAGACAATGAGATAAACATGAATACAGTATTACTCAGCAGCGTTCCAATT
TTCAACCCTCGTATTTATTAAAAAAAGGAATTTTTAATATATTTTTCTCCTTGTTGAAATA
TTTTAGTAACTGTTAATCGATATAGAGTGGAGTAGTGACGCTTTATTTCGGtTCATTCTCG
AAACAAAAATATAATAGTCCACTGAACTCTCTTAAATTGTTTTTACAACCTTCAACTGCCA
CAGACGTAATCCCTCACGTTATTTTGAGCTGACAACGTGTTGAATTGAGTGTGTTCCGAAT
TCTAAATAAGCATGTATATATTTACGTCTCATGCAAGTAATATATGTTTAACTGATGACGT
CACTTGGTGACCACTGATTTAGTTCCTTTGTCATAATTGCAGTTTCTGTTGTCACGGGGAC
GGTGGGGAAGCCAGGTTCCTCCTGTCACGCTGAATATCCCGTTCGAATCCCCCACATGGGT
ACAAAGTGTGATGCCTATTTCTGGTGTCCCCCACCGTGATATTGCTGGAATAAGTGGCTTA
ATACCATATACACTCACTCTATTGTCACACTACTGCCACCGGCTCACACCTCTGATGCTTC
TGTTCTATCCAG
```

Domain g(1)

```
GTCGCGCTGCTGACAGTGCGCACTCTGCCAACATTGCTGGCTCTGGGGTGAGGAAGGACGT
CACGACCCTCACTGTGTCTGAGACCGAGAACCTAAGACAGGCTCTTCAAGGTGTCATCGAT
GATACTGGTCCCAATGGTTACCAAGCAATAGCATCCTTCCACGGAAGTCCTCCAATGTGCG
AGATGAACGGCCGCAAGGTTGCCTGTTGTGCTCACG
```

Intron g(1)

```
GTAATTAATGGATGTGAAGTCAATGTCCGAGGGTATAATAAGGATTTAAATACTTCAGTCG
TGTAATACTGTATGACATGTGTATTGGATGGTGTAGGTATTACAGGTTATAAGGCCAGTGT
GTGTTGGGACGGTTACTTTCCTGCACTAGTAATAAGCATTGTATTTAGCTAGCTTTTATCA
TATAACTTTAGTTTCAGGTTTGtGGCAATTGAAATCGAAATTTTCTTTCATTTCAAGGTTA
TCGCACTCGTGTGTNAGAATAGTTACTATGCTGCATTGAGAATAACACTATAGTAATAAAG
CATATCATACAGTAAGAATAACACTATAGTAATAAAGTATATCATNCAGTAAGAATGTCAT
TGTATGATAAATAGGTTATCACACTCGTGTGTTTTAGAATGGTTACTATCCCAGGAATAAC
CACTATGTATTACATGTATATTGGGCAGTGTAAGTAGTAGCATTGTATATTAAATCAGTAT
ATCGTGCTTCAAAACACCAGGATATATGGGGTATACAGTGGGCAGTGTAAGTAGCAACATT
GTATATTAAATCAGTATATCGTACTTCAAAACACCAGGATTATGGGGTATACAGTGGGCAG
TGTAAGTAGTAGCATTGTATATTAAATCAGTATATCGTACTTCAAAACACCAGGATATAAT
TCAGTATATCGTGCTTCAAAACACCAGGATATAATTCAGTATATCGTGCTTCAAAACACCA
GGATATATGGGATATACAGTGCGGGTTTGCATACAACCTCCACCCTTTACAG
```

Domain g(2)

```
GTATGGcCTCCTTCCcACACTGGCACAGACTGTATGTGAAGCAGATGGAAGATGCCCTGGC
TGACCACGGGTCACATATCGGCATCCCTTACTGGGACTGGACAACTGCCTTCACAGAGTTA
CCCGCCCTTGTCACAGACTCCGAGAACAATCCCTTCCATGAG
```

Intron g(2)

GTCAGTTTAGTCTCCTGTCTGAGCTAACGATACCAATTTCCTATTTTCGAGAACCACGATG
ACGAGAAAACAAGCAATATAGATATAGATGCAGTATAGATCAAGTTAATGAATTCATTGCT
ATATGTTTGCTTGTAATAAACTTTAAGAAAACGAGAGCATGCACACAAATGAAACAAACAA
TTATGTGTTTGATAGGAATATGATATATGTATTTGGGGGCTGACGTGAGCAGGGTTGAAGG
GACAGTTTACATTGTCAGTAACACTGGGAGTATTCTTTGATCCACAATATATAGTTTCATT
GTGTTCAGCAGTTACAACTAACATTATATCATACATTACGTCGtAACATGCTTCTTTTGTC
CTCTTTTGCCAG

Domain g(3)

GGTCGCATTGATCATCTCGGTGTAACCACGTCACGTTCCCCCAGAGACATGCTGTTTAACG
ACCCAGAGCAAGGATCAGAGTCGTTCTTCTATAGACAAGTCCTCCTGGCTTTGGAGCAGAC
TGACTACTGCCAGTTCGAAGTCCAGTTTGAGCTGACCCACAACGCCATTCACTCCTGGACA
GGTGGACGTAGCCCTTACGGAATGTCGACCCTCGAGTTCACAGCCTACGATCCTCTCTTCT
GGCTTCACCACTCCAACACCGACAGAATCTGGGCTGTCTGGCAAGCACTGCAGAAATACCG
AGGACTCCCATACAACGAAGCACACTGTGAAATCCAGGTTCTGAAACAGCCCTTGAGGCCA
TTCAACGATGACATCAACCACAATCCAATCACCAAGACTAATGCCAGGCCTATCGATTCAT
TTGATTATGAGAGGtTTAACTATCAGTATGACACCCTTAGCTTCCATGGTAAGAGCATCCC
TGAACTGAATGACCTGCTCGAGGAAAGAAAAAGAGAAGAGAGAACATTTGCTGcCTTCCTT
CTTCGTGGAATCGGTTGCAGTGCTGATGTCGTCTTTGACATCTGCCGgCCCAATGGTGACT
GTGTCTTTGCAGGAACCTTTGCTGTGCTGGGAGGGGAGCTaGAAATGCCTTGGTCCTTCGA
CAGACTGTTCCGCTATGACATCACCAGAGTCATGAATCAGCTCCATCTCCAGTATGATTCA
GATTTCAGTTTCAGGGTGAAGCTTGTTGCCACCAATGGCACTGAGCTTTCATCAGACCTtC
TCAAGTCACCAACAATTGAACATGAACTTGG

Intron g/h

GTATGTTATCTTATCATCAAATGTGTGATCAGATACTGGAGACGTTTTCATATTAACTTGG
TCAGCATTAGTTGATGATTTTGGTGCGATGTTGACGACAAGGAGTCAAGCATTAACACATT
CAACACATCTTTAATCTGATATGAGAAGGGAATAAATTGATCCAGTATTGATGATTGAAGT
TAGATTAACAGTGAAAGATATACCAGTTTTGATAATCGTATAAAACAGTAGCAGAATTGTA
TCGTGAAAACTAAATGTGGGAAGGCGAACGCCAAGCAGATTTTAGATTACGATCGTGTGCT
AGAATAATTCACAATAACCCAGACGTCGGAAATGTGGTTGTCTATGGCAATGGTTACGATT
AATTGCTAACATGCACGATTTACCTATTTCAG

Domain h

AGCCCACAGAGGACCAGTTGAAGAAACAGAAGTCACTCGCCAACATACTGACGGCAATGCA
CACTTTCATCGTAAGGAAGTTGATTCGCTGTCCCTGGATGAAGCAAACAACTTGAAGAATG
CCCTTTACAAGCTACAGAACGACCACAGTCTAACGGGATACGAAGCAATCTCTGGTTACCA
TGGATACCCCAATCTGTGTCCGGAAGAAGGCGATGACAAAATACCCCTGCTGCGTCCCCGG
ATGGGCATCTTTCCTTACTGGCACAGACTCTTGACCATTCAACTGGAAAGAGCTCTTGAGC
ACAATGGTGCACTGCTTGGTGTTCCTTACTGGGACTGGAACAAGGACCTGTCGTCACTGCC
GGCGTTCTTCTCCGACTCCAGCAACAACAATCCCTACTTCAAGTACCACATCGCCGGTGTT
GGTCACGACACCGTCAGAGAGCCAACTAGTCTTATATATAACCAGCCCCAAATCCATGGTT
ATGATTATCTCTATTACCTAGCATTGACCACGCTTGAAGAAAACAATTACTGGGACTTTGA
GGTTCAGTATGAGATCCTCCACAACGCCGTCCACTCCTGGCTTGGAGGATCCCAGAAGTAT
TCCATGTCTACCCTGGAGTATTCGGCCTTTGACCCTGTCTTTATGATCCTTCACTCGGGTC
TAGACAGACTTTGGATCATCTGGCAAGAACTTCAGAAGATCAGGAGAAAGCCCTACAACTT
CGCTAAATGTGCTTATCATATGATGGAAGAGCCACTGGCGCCCTTCAGCTATCCATCTATC
AACCAGGACGAGTTCACCCGTGCCAACTCCAAGCCTTCTACAGTTTTTGACAGCCATAAGT

TCGGCTACCATTACGATAACCTGAATGTTAGAGGTCACAGCATCCAAGAACTCAACACAAT
CATCAATGACTTGAGAAACACAGACAGAATCTACGCAGGATTTGTTTTGTCAGGCATCGGT
ACGTCTGCTAGTGTCAAGATCTATCTCCGAACAGATGACAATGACGAAGAAGTTGGAACTT
TCACTGTCCTGGGAGGAGAGAGGGAAATGCCATGGGCCTACGAGCGAGTTTTCAAGTATGA
CATCACAGAGGTTGCAGATAGACTTAAAATTAAGTTATGGGGACACCCTTTAACTTCCGGA
ACTGGAGATCACATCCTTACGAATGGAATCGGTGGTAAACAAGAGCCTACCCAAATCCTTT
CATCATCTACAGACCTGCCAATCATGACTACGATGTTCTTGTTATCCCAGTANGGAAGAAA
CCTTCACATCCCTCCCAAAGTTGTCGTCAAGAAAGGCACCCGCATCGAGTTCCACCCAGTC
GATGATTCAGTTACGAGACCAGTTGTTGATCTTGGAAGCTACACTGCACTCTTCAACTGTG
TGGTACCACCGTTCACATACCACGGATTCGAACTGAACCACGTCTATTCTGTCAAGCCTGG
TGACTACTATGTTACTGGACCCACGAGAGACCTTTGCCAGAATGCAGATGTCAGGATTCAT
ATCCATGTTGAGGATGAGTAA

3´UTR

CGCAACAGGT

Intron UTR

GAGATAAGAAACCCTTCTAACAGTAATACGACACCACATTACAGCTTAAACATGATTGCCA
TCGATGTTTTCATGTGTAGTATACGCTTTTCAGTTCTACATAATTTTGTTTTTCAAATCAA
GTTTAGCAAATGAATCTATCACTGGAAAATAGGGTAGGGTAGCCAAGTGGTTAAAGCGGTC
ACTGATCACGCCAAAGACGAGTGTCCTAACCTGCATGGGTACAAAAGTGAAGACCATTGCT
GGTGTCTACCGCCGTAATATTGTTTTTAGTATTGCTAAAACTTATACTCACCCATGCGCTG
TAAAAGTGGAATAATAATCATATTTCAACAAAAGCACAAAACCATTTCATTTTCATGAAAG
CCTCTTGTTCACCTGAAAGACGCAAGAGAACAATAGTTCCTAACATTATTTTCAGACATTG
GAAATGTCCTGCACGTGTAAACCATATATCCTTTGAAATTTTTACGACTGCATCGTATACA
ATTTATGATATAAATTTAAAACTTTAT

3`UTR

TTCTTGGTCTCCACATATTCACATATCAGCACCAAATGGTTTCGAAGGACATTGGCGTTCT
TCTCTGGCAATGCATTTCAATACAACATTGAAAATGACTTCAGCATATCAGTGTGCTTCGA
ACGTGTTCCGGAAGTACTCAAATGTGCTATGACTGAATTATTGTACATACATAACTTATTG
ATGTTCAATAAATAAATGTTGAAACGAAAAAAAAAAAAAAAAAAAAAA

Figure 7

Derived primary structure of HtH2

Domain b

HRLFVTQVEDALIRRGSPIGVPYWDWTQPMAHLPGLADNATYRDPISGDSRHNPFHDVEVA
FENGRTERHPDSRLFEQPLFGKHTRLFDSIVYAFEQEDFCDFEVQFEMTHNNIHAWIGGGE
KYSMSSLHYTAFDPIFYLRHSNTDRLWAIWQALQIRRNRPYKAHCAWSEERQPLKPFAFSS
PLNNNEKTYENSVPTNVYDYEGVLGYTYDDLNFGGMDLGQLEEYIQRQRQRDRTFAGFFLS
HIGTSANVEIIIDHGTLHTSVGTFAVLGGEKEMKWGFDRLYKYEITDELRQLNLRADDVFS
ISVKVTDVDGSELSSELIPSAAIIFERSH

Domain c

IDHQDPHHDTIIRKNVDNLTPEEINSLRRAMADLQSDKTAGGFQQIAAFHGEPKWCPSPDA
EKKFSCCVHGMAVFPHWHRLLTVQGENALRKHGCLGALPYWDWTRPLSHLPDLVLVSSRTT
PMPYSTVEARNPWYSGHIDTVGVDTTRSVRQELYEAPGFGHYTGVAKQVLLALEQDDFCDF
EVQFEIAHNFIHALVGGSEPYGMASLRYTTYDPIFYLHHSNTDRLWAIWQALQKYRGKPYN
SANCAIASMRKPLQPFGLTDEINPDDETRQHAVPFSVFDYKNNFNYEYDTLDFNGLSISQL
DRELSRRKSHDRVFAGFLLHGIQQSALVKFFVCKSDDDCDHYAGEFYILGDEAEMPWGYDR
LYKYEITEQLNALDLHIGDRFFIRYEAFDLHGTSLGSNIFPKPSVIHDEGA

Domain d

GHHQADEYDEVVTAASHIRKNLKDLSKGEVESLRSAFLQLQNDGVYENIAKFHGKPGLCDD
NGRKVACCVHGMPTFPQWHRLYVLQVENALLERGSAVSVPYWDWTETFTELPSLIAEATYF
NSRQQTFDPNPFFRGKISFENAVTTRDPQPELYVNRYYYQNVMLVFEQDNYCDFEIQFEMV
HNVLHAWLGGRATYSISSLDYSAFDPVFFLHHANTDRLWAIWQELQRYRKKPYNEADCAIN
LMRKPLHPFDNSDLNHDPVTFKYSKPTDGFDYQNNFGYKYDNLEFNHFSIPRLEEIIRIRQ
RQDRVFAGFLLHNIGTSATVEIFVCVPTTSGEQNCENKAGTFAVLGGETEMAFHFDRLYRF
DISETLRDLGIQLDSHDFDLSIKIQGVNGSYLDPHILPEPSLIFVPGSS

Domain e

SFLRPDGHSDDILVRKEVNSLTTRETASLIHALKSMQEDHSPDGFQAIASFHALPPLCPSP
SAAHRYACCVHGMATFPQWHRLYTVQFQDALRRHGATVGVPYWDWLRPQSHLPELVTMETY
HDIWSNRDFPNPFYQANIEFEGENITTEREVIADKLFVKGGHVFDKLVLQTSHPSAEQENY
CDFEIQFEILHNGVHTWVGGSRTYSIGHLHYAFYDPLFYLHHFQTDRIWAIWQELQEQRGL
SGDEAHCALEQMREPLKPFSFGAPYNWNQLTQDFSRPEDTFDYRKFGYEYDNLEFLGMSVA
ELDQYIIEHQENDRVFAGFLLSGFGGSASVNFQVCRADSTCQDAGYFTVLGGSAEMAWAFD
RLYKYDITETLEKMHLRYDDDFTISVSLTANNGTVLSSSLIPTPSVIFQRGH

Domain f

RDINTRSMSPNRVRRELSDLSARDLSSLKSALRDLQEDDGPNGYQALAAFHGLPAGCHDSR
GNEIACCIHGMPTFPQWHRLYTLQLEMALRRHGSSVAIPYWDWTKPISELPSLFTSPEYYD
PWHDAVVNNPFSKGFVKFANTYTVRDPQEMLFQLCEHGESILYEQTLLALEQTDYCDFEVQ
FEVLHNVIHYLVGGRQTYALSSLHYASYDPFFFIHHSFVDKMWVVWQALQKRRKLPYKRAD
CAVNLMTKPMRPFDSDMNQNPFTKMHAVPNTLYDYETLYYSYDNLEIGGRNLDQLQAEIDR

SRSHDRVFAGFLLRGIGTSADVRFWICRNENDCHRGGIIFILGGAKEMPWSFDRNFKFDIT
HVLENAGISPEDVFDAEEPFYIKVEIHAVNKTMIPSSVIPAPTIIYSPGE

Domain g

GRAADSAHSANIAGSGVRKDVTTLTVSETENLRQALQGVIDDTGPNGYQAIASFHGSPPMC
EMNGRKVACCAHGMASFPHWHRLYVKQMEDALADHGSHIGIPYWDWTTAFTELPALVTDSE
NNPFHEGRIDHLGVTTSRSPRDMLFNDPEQGSESFFYRQVLLALEQTDYCQFEVQFELTHN
AIHSWTGGRSPYGMSTLEFTAYDPLFWLHHSNTDRIWAVWQALQKYRGLPYNEAHCEIQVL
KQPLRPFNDDINHNPITKTNARPIDSFDYERFNYQYDTLSFHGKSIPELNDLLEERKREER
TFAAFLLRGIGCSADVVFDICRPNGDCVFAGTFAVLGGELEMPWSFDRLFRYDITRVMNQL
HLQYDSDFSFRVKLVATNGTELSSDLLKSPTIEHEL

Domain h

GAHRGPVEETEVTRQHTDGNAHFHRKEVDSLSLDEANNLKNALYKLQNDHSLTGYEAISGY
HGYPNLCPEEGDDKIPLLRPRMGIFPYWHRLLTIQLERALEHNGALLGVPYWDWNKDLSSL
PAFFSDSSNNNPYFKYHIAGVGHDTVREPTSLIYNQPQIHGYDYLYYLALTTLEENNYWDF
EVQYEILHNAVHSWLGGSQKYSMSTLEYSAFDPVFMILHSGLDRLWIIWQELQKIRRKPYN
FAKCAYHMMEEPLAPFSYPSINQDEFTRANSKPSTVFDSHKFGYHYDNLNVRGHSIQELNT
IINDLRNTDRIYAGFVLSGIGTSASVKIYLRTDDNDEEVGTFTVLGGEREMPWAYERVFKY
DITEVADRLKIKLWGHPLTSGTGDHILTNGIGGKQEPTQILSSSTDLPIMTTMFLLSQXGR
NLHIPPKVVVKKGTRIEFHPVDDSVTRPVVDLGSYTALFNCVVPPFTYHGFELNHVYSVKP
GDYYVTGPTRDLCQNADVRIHIHVEDE

Figure 8

KLH1 cDNA sequence and intron structure

Domain b

GGCCTACCGTACTGGGACTGGACTGAACCCATGACACACATTCCGGGTCTGGCAGGAAACA
AAACTTATGTGGATTCTCATGGTGCATCCCACACAAATCCTTTTCATAGTTCAGTGATTGC
ATTTGAAGAAAATGCTCCCCACACCAAAAGACAAATAGATCAAAGACTCTTTAAACCCGCT
ACCTTTGGACACCACACAGACCTGTTCAACCAGATTTTGTATGCCTTTGAACAAGAAGATT
ACTGTGACTTTGAAGTCCAATTTGAGATTACCCATAACACGATTCACGCTTGGACAGGAGG
AAGCGAACATTTCTCAATGTCGTCCCTACATTACACAGCTTTCGATCCTTTGTTTTACTTT
CACCATTCTAACGTTGATCGTCTTTGGGCCGTTTGGCAAGCCTTACAGATGAGACGGCATA
AACCCTACAGGGCCCACTGCGCCATATCTCTGGAACATATGCATCTGAAACCATTCGCCTT
TTCATCTCCCCTTAACAATAACGAAAAGACTCATGCCAATGCCATGCCAAACAAGATCTAC
GACTATGAAAATGTCCTCCATTACACATACGAAGATTTAACATTTGGAGGCATCTCTCTGG
AAAACATAGAAAAGATGATCCACGAAAACCAGCAAGAAGACAGAATATATGCCGGTTTTCT
CCTGGCTGGCATACGTACTTCAGCAAATGTTGATATCTTCATTAAAACTACCGATTCCGTG
CAACATAAGGCTGGAACATTTGCAGTGCTCGGTGGAAGCAAGGAAATGAAGTGGGGATTTG
ATCGCGTTTTCAAGTTTGACATCACGCACGTTTTGAAAGATCTCGATCTCACTGCTGATGG
CGATTTCGAAGTTACTGTTGACATCACTGAAGTCGATGGAACTAAACTTGCATCCAGTCTT
ATTCCACATGCTTCTGTCATTCGTGAGCATGCACGTGGTAAGCTGAATAGAG

Intron b/c

GTTTTGTAATAATTATGTAGAATTCTTTACCTCAGAATAAGATGAGGTCACATGGGTTTTG
CAAAACTATTACGTTCGAATTAATATTAATAATACCGGACCCTCCACTGGTACATATTTAT
CTTTATAACGATAATAGCGATGATGATGATGATGATGATGATGATGATGATGATgATAATg
ATgATGCCGGTATTGCACGTAATCCAGCCGAcTTAGATGACACCCTAAGGGTGCAGAAAGT
ATAaCAATTAGATTGCGTTtGCATCTGTGTATGCGTGTGCTTTAaCCAAAAGTCAAAATAA
AAGTGCAAACCCTTAGTTTATTCATTTGATAGAGCCTTTTACGATAAGAACAATGTAATAA
ATTAGAACATAACTGAAACCTCCGAAAGAAGGCCTGTTTGTCAAGAGAGGTATCGACATGA
TTGACTTATAAACCTGTGCTTCTATATTTTGGAACTGTCCACTTTCTTGTTGTGTGTACTG
TAATCACATCGCACTATGGCTGCAAGACGTGTACGAGTACACTATATACTTACCTAATGAC
CAACCACAAGGCTGGCTTTGTTAATATTGTTATTTCACAGAAATAAACACAGAATTCCAGC
ATTTGGCTGGTGTATTTAGCAAAACACCGATATGACACTCATGTTTTATTACATTTTTTC
AG

Domain c

TTAAATTTGACAAAGTGCCAAGGAGTCGTCTTATTCGAAAAAATGTAGACCGTTTGAGCCC
CGAGGAGATGAATGAACTTCGTAAAGCCCTAGCCTTACTGAAAGAGGACAAAAGTGCCGGT
GGATTTCAGCAGCTTGGTGCATTCCATGGGGAGCCAAAATGGTGTCCTAGTCCCGAAGCAT
CTAAAAAATTTGCCTGCTGTGTTCACGGCATGTCTGTGTTCCCTCACTGGCATCGACTGTT
GACGGTTCAGAGTGAAAATGCTTTGAGACGACATGGCTACGATGGAGCTTTGCCGTACTGG
GATTGGACCTCTCCTCTTAATCACCTTCCCGAACTGGCAGATCATGAGAAGTACGTCGACC
CTGAAGATGGGGTAGAGAAGCATAACCCTTGGTTCGATGGTCATATAGATACAGTCGACAA
AACAACAACAAGAAGTGTTCAGAATAAACTCTTCGAACAGCCTGAGTTTGGTCATTATACA
AGCATTGCCAAACAAGTACTGCTAGCGTTGGAACAGGACAATTTCTGTGACTTTGAAATCC
AATATGAGATTGCCCATAACTACATCCATGCACTTGTAGGAGGCGCTCAGCCTTATGGTAT
GGCATCGCTTCGCTACACTGCTTTTGATCCACTATTCTACTTGCATCACTCTAATACAGAT
CGTATATGGGCAATATGGCAGGCTTTACAGAAGTACAGAGGAAAACCGTACAACGTTGCTA
ACTGTGCTGTTACATCGATGAGAGAACCTTTGCAACCATTTGGCCTCTCTGCCAATATCAA
CACAGACCATGTAACCAAGGAGCATTCAGTGCCATTCAACGTTTTTGATTACAAGACCAAT

TTCAATTATGAATATGACACTTTGGAATTTAACGGTCTCTCAATCTCTCAGTTGAATAAAA
AGCTCGAAGCGATAAAGAGCCAAGACAGGTTCTTTGCAGGCTTCCTGTTATCTGGTTTCAA
GAAATCATCTCTTGTTAAATTCAATATTTGCACCGATAGCAGCAACTGTCACCCCGCTGGA
GAGTTTTACCTTCTGGGTGATGAAAACGAGATGCCATGGGCATACGATAGAGTCTTCAAAT
ATGACATAACCGAAAAACTCCACGATCTAAAGCTGCATGCAGAAGACCACTTCTACATTGA
CTATGAAGTATTTGACCTTAAACCAGCAAGCCTGGGAAAAGATTTGTTCAAGCAGCCTTCA
GTCATTCATGAACCAAGAATAG

Intron c/d

GTACTTGTTATATGTTTCGAATATTGCCGATACCTTCAATATATATACTTTATCAAAGTAA
TTGATTAATCTGAAGTAATTTTCCTTTCCAGTAGAGATTCAGTTGATACAACAAGAATTCG
CCCTGTTGTATGTCACTTTATTTTCATCAAACGATTCGAAGTGAGCTGTCCATGCCACAAT
GGGGTCTCTGTAACTTTCTCGTATGGGGTATAGATTATATAGACGTGGCAGACCTTACGTA
TAACTAATATTTGTGTAATGTCGTTTCAG

Domain d

GTCACCATGAAGGCGAAGTATATCAAGCTGAAGTAACTTCTGCCAACCGTATTCGAAAAAA
CATTGAAAATCTGAGCCTTGGTGAACTCGAAAGTCTGAGAGCTGCCTTCCTGGAAATTGAA
AACGATGGAACTTACGAATCAATAGCTAAATTCCATGGTAGCCCTGGTTTGTGCCAGTTAA
ATGGTAACCCCATCTCTTGTTGTGTCCATGGCATGCCAACTTTCCCTCACTGGCACAGACT
GTACGTGGTTGTCGTTGAGAATGCCCTCCTGAAAAAAGGATCATCTGTAGCTGTTCCCTAT
TGGGACTGGACAAAACGAATCGAACATTTACCTCACCTGATTTCAGACGCCACTTACTACA
ATTCCAGGCAACATCACTATGAGACAAACCCATTCCATCATGGCAAAATCACACACGAGAA
TGAAATCACTACTAGGGATCCCAAGGACAGCCTCTTCCATTCAGACTACTTTTACGAGCAG
GTCCTTTACGCCTTGGAGCAGGATAACTTCTGTGATTTCGAGATTCAGTTGGAGATATTAC
ACAATGCATTGCATTCTTTACTTGGTGGCAAAGGTAAATATTCCATGTCAAACCTTGATTA
CGCTGCTTTTGATCCTGTGTTCTTCCTTCATCACGCAACGACTGACAGAATCTGGGCAATC
TGGCAAGACCTTCAGAGGTTCCGAAAACGGCCATACCGAGAAGCGAATTGCGCTATCCAAT
TGATGCACACGCCACTCCAGCCGTTTGATAAGAGCGACAACAATGACGAGGCAACGAAAAC
GCATGCCACTCCACATGATGGTTTTGAATATCAAAACAGCTTTGGTTATGCTTACGATAAT
CTGGAACTGAATCACTACTCGATTCCTCAGCTTGATCACATGCTGCAAGAAAGAAAAAGGC
ATGACAGAGTATTCGCTGGCTTCCTCcTTCACAATATTGGAACatCTGCCGATGGCCATGT
ATTTGTATGTCTCCCAACTGGGGAACACACGAAGGACTGCAGTCATGAGGCTGGTATGTTC
TCCATCTTAGGCGGTCAAACGGAgATGTCCTTTGTATTTGACAGACTTTACAAACTTGACA
TAACTAAAGCCTTGAAAAAGAACGGTGTGCACCTGCAAGGGGATTTCGATCTGGAAATTGA
GATTACGGCTGTGAATGGATCTCATCTAGACAGTCATGTCATCCACTCTCCCACTATACTG
TTTGAGGCCGGAACAG

Intron d/e

GTAACTATTTTGTCACTGTAACCAACAACTGCAGTCTATTTTGCAATTACGATAATAACAA
TTTTTGAAATATATCTTTATTAAAGCAAAGGTTTCTAGAGACAAACAGCCGGCTCTAATTA
TTTTTTCGAACTTACGCTTGAGTAAAGATCTGCAAATGGCAACCCTACCTATACTATTAAA
AATATAATGTTACATTCGTATCTGAATGTTTAATAAATCACTTCATATTCTGTTGCAG

Domain e

ATTCTGCCCACACAGATGATGGACACACTGAACCAGTGATGATTCGCAAAGATATCACACA
ATTGGACAAGCGTCAACAACTGTCACTGGTGAAAGCCCTCGAGTCCATGAAAGCCGACCAT
TCATCTGATGGGTTCCAGGCAATCGCTTCCTTCCATGCTCTTCCTCCTCTTTGTCCATCAC
CAGCTGCTTCAAAGAGGTTTGCGTGCTGCGTCCATGGCATGCCAACCTTCCCGCAATG

Figure 9

Derived primary structure of KLH1

Domain b

```
GLPYWDWTEPMTHIPGLAGNKTYVDSHGASHTNPFHSSVIAFEENAPHTKRQIDQRLFKPA
TFGHHTDLFNQILYAFEQEDYCDFEVQFEITHNTIHAWTGGSEHFSMSSLHYTAFDPLFYF
HHSNVDRLWAVWQALQMRRHKPYRAHCAISLEHMHLKPFAFSSPLNNNEKTHANAMPNKIY
DYENVLHYTYEDLTFGGISLENIEKMIHENQQEDRIYAGFLLAGIRTSANVDIFIKTTDSV
QHKAGTFAVLGGSKEMKWGFDRVFKFDITHVLKDLDLTADGDFEVTVDITEVDGTKLASSL
IPHASVIREHARGKLNR
```

Domain c

```
VKFDKVPRSRLIRKNVDRLSPEEMNELRKALALLKEDKSAGGFQQLGAFHGEPKWCPSPEA
SKKFACCVHGMSVFPHWHRLLTVQSENALRRHGYDGALPYWDWTSPLNHLPELADHEKYVD
PEDGVEKHNPWFDGHIDTVDKTTTRSVQNKLFEQPEFGHYTSIAKQVLLALEQDNFCDFEI
QYEIAHNYIHALVGGAQPYGMASLRYTAFDPLFYLHHSNTDRIWAIWQALQKYRGKPYNVA
NCAVTSMREPLQPFGLSANINTDHVTKEHSVPFNVFDYKTNFNYEYDTLEFNGLSISQLNK
KLEAIKSQDRFFAGFLLSGFKKSSLVKFNICTDSSNCHPAGEFYLLGDENEMPWAYDRVFK
YDITEKLHDLKLHAEDHFYIDYEVFDLKPASLGKDLFKQPSVIHEPRI
```

Domain d

```
GHHEGEVYQAEVTSANRIRKNIENLSLGELESLRAAFLEIENDGTYESIAKFHGSPGLCQL
NGNPISCCVHGMPTFPHWHRLYVVVVENALLKKGSSVAVPYWDWTKRIEHLPHLISDATYY
NSRQHHYETNPFHHGKITHENEITTRDPKDSLFHSDYFYEQVLYALEQDNFCDFEIQLEIL
HNALHSLLGGKGKYSMSNLDYAAFDPVFFLHHATTDRIWAIWQDLQRFRKRPYREANCAIQ
LMHTPLQPFDKSDNNDEATKTHATPHDGFEYQNSFGYAYDNLELNHYSIPQLDHMLQERKR
HDRVFAGFLLHNIGTSADGHVFVCLPTGEHTKDCSHEAGMFSILGGQTEMSFVFDRLYKLD
ITKALKKNGVHLQGDFDLEIEITAVNGSHLDSHVIHSPTILFEAG
```

Domain e

```
DSAHTDDGHTEPVMIRKDITQLDKRQQLSLVKALESMKADHSSDGFQAIASFHALPPLCPS
PAASKRFACCVHGMPTFPQWHRLYTVQFQDSLRKHGAVVGLPYWDWTLPR
```

Figure 10

KLH2 cDNA sequence and intron structure

Domain b

GGCCTGCCCTACTGGGATTGGACCATGCCAATGAGTCATTTGCCAGAACTGGCTACAAGTG
AGACCTACCTCGATCCAGTTACTGGGGAAACTAAAAACAACCCTTTCCATCACGCCCAAGT
GGCGTTTGAAAATGGTGTAACAAGCAGGAATCCTGATGCCAAACTTTTTATGAAACCAACT
TACGGAGACCACACTTACCTCTTCGACAGCATGATCTACGCATTTGAGCAGGAAGACTTCT
GCGACTTTGAAGTCCAATATGAGCTCACGCATAATGCAATACATGCATGGGTTGGAGGCAG
TGAAAAGTATTCAATGTCTTCTCTTCACtacacTGCTTTTGATCCTATATTTTACCTCCAT
CACTCAAATGTTGATCGTCTCTGGGCCATTTGGCAAGCTCTTCAAATCAGGAGAGGCAAGT
CTTACAAGGCCCACTGCGCCTCGTCTCAAGAAAGAGAACCATTAAAGCCTTTTGCATTCAG
TTCCCCACTGAACAACAACGAGAAAACGTACCACAACTCTGTCCCCACTAACGTTTATGAC
TATGTGGGAGTTTTGCACTATCGATATGATGACCTTCAGTTTGGCGGTATGACCATGTCAG
AACTTGAGGAATATATTCACAAGCAGACACAACATGATAGAACCTTTGCAGGATTCTTCCT
TTCATATATTGGAACATCAGCAAGCGTAGATATCTTCATCAATCGAGAAGGTCATGATAAA
TACAAAGTGGGAAGTTTTGTAGTACTTGGTGGATCCAAAGAAATGAAATGGGGCTTTGATA
GAATGTACAAGTATGAGATCACTGAGGCTCTGAAGACGCTGAATGTTGCAGTGGATGATGG
GTTCAGCATTACTGTTGAGATCACCGATGTTGATGGATCTCCCCCATCTGCAGATCTCATT
CCACCTCCTGCTATAATCtTTGaACGTGGTCaTG

Intron 2b/c

AGGTATTTAAAAAAGTAATAAAACCaTATTTTCGAATGCGCTTTATGAAATATCGTGTGAC
TGGTTCTTTAGTTTACATGGAGTGTAACAACATGCTCCATCAGTTGACATATACTGCTCAC
ACAAAGTAAGGGATATTTGATAATGATAACAAATATAATCAAAGCGGTTATACTATCAAGA
CTTATTCACATAATTACAGGTGAAGGGAGGTGTGATCGTGTTCACTGATCAGGTTGAGGCC
AGAGAAGTCCCAGTTTGAGTCTTGCAGAAGATGATGTTTAGGCATGGGGTCGAATCACCAA
AATCACATGACTTCAATAACGGGTTGGACCACCTCGAGCGACgATGCAAGCAGTAGAGCGT
CTACGCATGCTCCTGATAAGGCGACCAATCTGTTCCTGGGGAATCAGtCGCCACTCCTCTT
GTAGTGCCACGCTCATTTCTGCTACGGTCCTGGGTACCTGCTATCGGgTCTTGATCCGTAT
CCCAAGGATGTCCCACACATGTTCAAgGTGAGAGGTCGGGGAACATCGCTGGCCACGGTaA
GGtCTGAATTTGATGCCGTTGAAAGTGAGCTCTGACAACcTGAGCATGGtGAGCTCTGACG
TTGTCGTCCTGAAAGATGAATcCAGCTcCaTGaCAGCGAGCAAaGGGCAGGACGTGTTGGT
CAATGCAGTTGTCTCTGCAGTACACACCTGTCACTCGCCACTCACAAGCGTGTAGATCTGT
ACGACCAGTCATGGAGATCCCAGCCCACATCATAACGGACCCCTATCCATACCGATCATGA
GCCACCATAGCAGCGTCTTGATGACGTTCTCCCTGTCGCCTCGACATCCTcACACGGCCAA
AAGGAACGTGGACTCGTCACTGAACATGACATTAGCCAACCTGGCACTTGTCCACCGCTGA
TGTTGGCGAGACCATTCCAGTCGAGCTCTTCGGTGTCTGGCTTTCATCGATAACACGACGT
AAGGTCTGCGGGCGTGCAAGACGGCTCTATGCAGGCGATTTCGGATTGTCTGGGTGCTAAC
TCTGATCCCAGGTGCCTGCTGAAGTTGATGCTGGATCTGTGTGGCATTGAGATGGCGATTC
CTTAGGACTGTGGAGATGATGAATCGATCTTGACTTATGGTGGTGACATTAGGACGTCGGG
TTCGTGTCCTATCCTGCACTCTTCCAGTTGTTCGGTGACGCTCTGGTACCCGGCTGATTAC
TGACTGAGAATATCCATCTGCCGTGCGACATGAGCCTGTGTTGGCCCAGCCTGAAGCATTG
CAATCGCCAGAGACGCTCTTCAAAAGTCATTCGACGCATGGtTTTCTGTTCACAAATGACA
GCGTAAAACAGtTTTTGGtGCTTTTATGCTTCCCAAGAGCATGAAAAACACGTTCTATgGG
TCGtGCACACCTTACATGACAAGtGtGAAAAGtGACTTGcACCCCCTTGtGtGTTCGGATG
CACACTCTGTTTACGTACTGATGCGATTTGGCGTCTAAACATGTTTTGGCGTCTAAACATG
TTTTCCTGCATGATTCATATACTATTTTGTCATATTCCTGGCATCAAACCAAACTACAGTG
AAATATATTTCAATATCCCCTACTTTGTGTGAGTAGTATAGATCACTGCAGACAACATATA

GACAAtGCAgtTaCaCCGTCAACAATCCCAGTCATTAATTATGATGaCaCTTCCACACATA
GTGTCAGTGATTGTAATTCAaCTGTACACACTTTTCCCGTGAACATTCAGGATCTATATGA
CTAAATATATAACATTAGTATACGTGCAGTTTTGTATCGCTACGACATTGTTGTAACTCTT
TGTTTAATCATTTaACAG

Domain c

CTGATGCCAAAGaCTTTGgCCATAGCAGAAAAATCAGgAAAGcCGTTGATTcTcTGACAGT
CgAAGAACAAAcTTCGTTGAGgCGAGcTATGgCAGATcTACAGGACGACAAAACATCAGGG
GGTTTCCAGCAGATTGCAGCATTCCACGGAGAACCAAAATGGTGTCCAAGCCCCgAAGCGG
AGAAAAAATTTGCATGCTGTGTTCATGGAATGGCTGTTTTCCCTCACTGGCACAGATTGCT
GACAGTTCAAGGAGAAAATGCTCTGAGGAAACATGGCTTTACTGGTGGACTGCCCTACTGG
GACTGGACTCGATCAATGAGCGCCCTTCCACATTTTGTTGCTGATCCTACTTACAATGATG
CTATTTCCAGCCAGGAAGAAGATAACCCATGGCATCATGGTCACATAGACTCTGTTGGGCA
TGATACTACAAGAGATGTGCGTGATGATCTTTATCAATCTCCTGGTTTCGGTCACTACACA
GATATTGCACAACAAGTCCTTCTGGCCTTTGAGCAGGACAGTTTCTGTGATTTTGAGGTAC
AATTTGAAATTGCCCATAATTTCATACATGCACTGATTGGTGGTAACGAACCATACAGTAT
GTCATCTTTGAGGTATACTACATACGATCCAATCTTCTTCTTGCACCACTCCAGTACAGAC
CGACTTTGGGCCATCTGGCAAGCAATCACTAGTGCGGCCGCCTGCAGGTCGACCATAAGGG
AGAGCTCCCAACGCgtTGGAtGCAATCT

Domain g

ATGGCTGTGTTTCCGCACTGGCACAGACTGTTTGTGAAACAGATGGAGGACGCACTTGCTG
CTCATGGAGCTCATATTGGCATACCATACTGGGATTGGACAAGTGCGTTTAGTCATCTGCC
CGCCCTAGTGACTGACCACGAGAACAATCCCTTCCACCAC

Intron g(2)

GTATGTGTCAAATCGTTTTAGGAACTGCCTTATCCATTTTACAATTACGAGTACAAAATGA
AAACGGAAACTGTGTGACCTCGAAAAGTGCAATCTTTAAAGGATGCAATGTACACAATAAA
ATGCTCCGATCAAAAGCGATGGCTAGAAATCATTTTCCCTCTAATTCCCTTTCACACAGCT
CGGTTCGTTTTAAGTAGGAACAAGTCTCTGCAAAAACATCACAAATAAAGAGAACACAGAA
AAAACCTCATTCTCGTTTCTGTATTCCGAAAATGAAATTTACAATTTCTTTCATTTATAG

Domain g

GGCCATATTGGTCATCTGAATGTGGATACATCTCGATCTCCAAGAGACATGCTGTTTAATG
ATCCTGAACAAGGCTCAGAATCATTCTTCTACAGACAGGTTCTCTTGACTCTAGAACAGAC
AGACTTCTGCCAATTTGAAGTTCAGTTTGAACTTACACACAATGCCATCCACTCTTGGACT
GGAGGACATACTCCATATGGAATGTCATCACTGGAATATACAGCATATGATCCACTCTTTT
ATCTCCACCATTCCAACACTGATCGTATCTGGGCCATCTGGCAGGCACTCCAGAAATATAG
AGGTCTTCCATACAACGCAGCTCACTGCGATATccaagttctgaaacaacctcTTAAACCA
TTCAGCGAGTCCAGGAATCCAAACCCAGTCACCAGAGCCAATTCTAGGGCCGTTGATTCAT
TTGATTATGAGAAATTCAATTATCAATATGACACACTTACCTTCCACGGACTTTCTATCCC
AGAACTTGATGCCATGCTTCAAGAGAGAAAGAAGGAAGAGAGAACATTTGCAGCCTTCCTG
TTGCACGGATTTGGCGCCAGTGCTGATGTTTCGTTTGATGTCTGCACACCTGATGGTCATT
GTGCCTTTGCTGGAACCTTCGCGGTACTTGGTGGGGAGCTTGAGATGCCCTGGTCCTTTGA
AAGATTGTTCCGTTACGATATCACAAAGGTTCTCAAGCAGATGAATCTTCACTATGATTCT
GAGTTCCACTTTGAGTTGAAGATTGTTGGCACAGATGGAACAGAACTGCCATCGGATCGTA
TCAAGAGCCCTACCATTGAACACCATGGAGGAG

Intron g/h

GTATGTTTTGAGATCCACATAATCTTCTACCCTGTCTCATTTCTAATGCTCTTCAATACAC
AATTTATATAGCCTTTGAGCTTCAGATGTATTACGGACAGGCATTACAGTATACATGTAAT
ATGGTTTTCTGCTATTTGCAAAAATTGTGTCCTATCTCTGTTCAGATCATCATGGCGGTGA
CACCTAG

Domain h

GTCACGATCACAGTGAACGTCACGATGGATTTTTCAGGAAGGAAGTCGGTTCCCTGTCCCT
GGATGAAGCCAATGACCTTAAAAATGCACTGTACAAGCTGCAGAATGATCAGGGTCCCAAT
GGATATGAATCAATAGCCGGTTACCATGGCTATCCATTCCTCTGCCCTGAACATGGTGAAG
ACCAGTACGCATGCTGTGTCCACGGAATGCCTGTATTTCCACATTGGCACAGACTTCATAC
AATCCAGTTTGAGAGAGCTCTCAAAGAACATGGTTCTCATTTGGGTCTGCCATACTGGGAC
TGGAC

Figure 11

Derived primary structure of KLH2

Domain b

GLPYWDWTMPMSHLPELATSETYLDPVTGETKNNPFHHAQVAFENGVTSRNPDAKLFMKPT
YGDHTYLFDSMIYAFEQEDFCDFEVQYELTHNAIHAWVGGSEKYSMSSLHYTAFDPIFYLH
HSNVDRLWAIWQALQIRRGKSYKAHCASSQEREPLKPFAFSSPLNNNEKTYHNSVPTNVYD
YVGVLHYRYDDLQFGGMTMSELEEYIHKQTQHDRTFAGFFLSYIGTSASVDIFINREGHDK
YKVGSFVVLGGSKEMKWGFDRMYKYEITEALKTLNVAVDDGFSITVEITDVDGSPPSADLI
PPPAIIFERGHA

Domain c

DAKDFGHSRKIRKAVDSLTVEEQTSLRRAMADLQDDKTSGGFQQIAAFHGEPKWCPSPEAE
KKFACCVHGMAVFPHWHRLLTVQGENALRKHGFTGGLPYWDWTRSMSALPHFVADPTYNDA
ISSQEEDNPWHHGHIDSVGHDTTRDVRDDLYQSPGFGHYTDIAQQVLLAFEQDSFCDFEVQ
FEIAHNFIHALIGGNEPYSMSSLRYTTYDPIFFLHHSSTDRLWAIWQALQKYRGKPYNTAN
CAIASMRKPLQPFGLDSVINPDDETREHSVPFRVFDYKNNFDYEYESLAFNGLSIAQLDRE
LQRRKSHDRVFAGFLLHEIGQSAKHNVSDCDHYAGEFYILGDEAEMPWRYDRVYKYEITQQ
LHDLDLHVGDNFFLKYEAFDLNGGSLGGSIFSQPSVIFEPAAGMF

Domain d

GSHQADEYREAVTSASHIRKNIRDLSEGEIESIRSAFLQIQKEGIYENIAKFHGKPGLCEH
DGHPVACCVHGMPTFPHWHRLYVLQVENALLERGSAVAVPYWDWTLPR

Domain g

MAVFPHWHRLFVKQMEDALAAHGAHIGIPYWDWTSAFSHLPALVTDHENNPFHHGHIGHLN
VDTSRSPRDMLFNDPEQGSESFFYRQVLLTLEQTDFCQFEVQFELTHNAIHSWTGGHTPYG
MSSLEYTAYDPLFYLHHSNTDRIWAIWQALQKYRGLPYNAAHCDIQVLKQPLKPFSESRNP
NPVTRANSRAVDSFDYEKFNYQYDTLTFHGLSIPELDAMLQERKKEERTFAAFLLHGFGAS
ADVSFDVCTPDGHCAFAGTFAVLGGELEMPWSFERLFRYDITKVLKQMNLHYDSEFHFELK
IVGTDGTELPSDRIKSPTIEHHGG

Domain h

GHDHSERHDGFFRKEVGSLSLDEANDLKNALYKLQNDQGPNGYESIAGYHGYPFLCPEHGE
DQYACCVHGMPVFPHWHRLHTIQFERALKEHGSHLGLPYWDW